

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 9, 1999, 10:26:01 ; Search time 23.43 seconds  
(without alignments)  
178,687 Million cell updates/sec

Title: US-08-486-814-17  
Perfect score: 1080  
Sequence: 1 MSRLDKSVNSALELLNEV.....FGLELIICGLEKQKCESGS 207

Scoring table: BLOSUM62

Searched: 162890 seqs, 20225328 residues

Database : A\_Geneseq\_34.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1080	100.0	335	1	R64808	tTA transactivator
2	1080	100.0	207	1	R89763	Wild type Tn-10 de
3	1080	100.0	207	1	W08296	Wild-type E.coli T
4	1080	100.0	651	1	W48357	Multi-chimeric tra
5	1080	100.0	651	1	W47593	Multi-chimeric tra
6	1080	100.0	207	1	W1326	Protein encoded by
7	1075	99.5	207	1	R48630	Sequence of the te
8	1075	99.5	297	1	R64809	tTAS transactivato
9	1075	99.5	297	1	R85324	tTAS transactivato
10	1075	99.5	1088	1	R88636	Plasmid pASK75 ope
11	1075	99.5	354	1	W21994	Tetracycline trans
12	1075	99.5	349	1	W54311	Plectestrin homolog
13	1073	99.4	335	1	R53323	tTA transactivator
14	1070	99.1	336	1	W08474	pret-tTAK encoded
15	1068	98.9	207	1	W08323	Mutated Tet repres
16	1067	98.8	207	1	W08317	Mutated Tet repres
17	1067	98.8	207	1	W08318	Mutated Tet repres
18	1067	98.8	207	1	W08300	Mutated Tet repres
19	1066	98.7	207	1	W08320	Mutated Tet repres
20	1066	98.7	207	1	W08304	Mutated Tet repres
21	1065	98.6	207	1	W08312	Mutated Tet repres
22	1063	98.4	207	1	W08306	Mutated Tet repres
23	1061	98.2	207	1	W08325	Mutated Tet repres
24	1061	98.2	207	1	W08326	Mutated Tet repres
25	1059	98.1	207	1	W08301	Mutated Tet repres
26	1058	98.0	335	1	R89761	Mutated Tn-10 deri
27	1058	98.0	207	1	R89764	Mutated Tn-10 deri
28	1058	98.0	335	1	W08293	Mutated reverse Te
29	1058	98.0	207	1	W08297	Mutated Tet repres
30	1058	98.0	207	1	W08321	Mutated Tet repres
31	1058	98.0	207	1	W08305	Mutated Tet repres
32	1058	98.0	335	1	W71323	Protein encoded by
33	1058	98.0	207	1	W71327	Protein encoded by
34	1057	97.9	207	1	W08303	Mutated Tet repres
35	1057	97.9	207	1	W08310	Mutated Tet repres
36	1057	97.9	207	1	W08313	Mutated Tet repres
37	1056	97.8	207	1	W08316	Mutated Tet repres
38	1056	97.8	207	1	W08319	Mutated Tet repres
39	1056	97.8	207	1	W08302	Mutated Tet repres
40	1056	97.8	207	1	W08308	Mutated Tet repres
41	1055	97.7	207	1	W08322	Mutated Tet repres
42	1055	97.7	207	1	W08324	Mutated Tet repres
43	1055	97.7	207	1	W08307	Mutated Tet repres

ALIGNMENTS

44 1054 97.6 207 1 W08309 Mutated Tet repres  
45 1054 97.6 207 1 W08311 Mutated Tet repres

RESULT 1  
R64808 R64808 standard; Protein; 335 AA.  
AC R64808:  
DT 17-JUL-1995 (first entry)  
DE tTA transactivator.  
KW tTA; transactivator; tetracycline-controllable transactivator;  
KW conditional inactivation; homologous recombination; gene expression;  
KW gene regulation; gene therapy; tetracycline-resistance; tetr;  
KW transgenic animal.  
OS Herpes simplex virus K12, KOS.  
PN W09429442-A.  
PD 22-DEC-1994.  
PF 14-JUN-1994; U06734.  
PR 14-JUN-1993; US-076327.  
PA (BADI ) BASF AG.  
PI Bujard H, Gossen M, Salfeld JG, Voss JW;  
DR WPI: 95-036472/05.  
DR N-PSDB; Q76264.  
PT Regulatory systems using tetracycline-controllable transactivator  
PT (tTA) - useful for conditional inactivation or modulation of  
PT gene expression in a host cell or animal.  
PS Disclosure: Page 46-48; 103pp; English.  
CC A 397 bp MuL/PokI fragment of pMSVP16 coding for the C-terminal 130  
CC amino acids of VP16 of HSV was blunted and inserted in pUHD14-1.  
CC The resulting plasmid, pUHD15-1, encoded a tetr-VP16 fusion protein,  
CC or tetracycline-controllable transactivator (tTA), whose sequence is  
CC given in Q76264, and the encoded protein sequence in R64808.  
SQ Sequence 335 AA;

Query Match 100.0%; Score 1080; DB 1; Length 335;  
Best Local Similarity 100.0%; Pred. No. 1.2e-112;  
Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSRLDKSVNSALELLNEVGIEGLTRKLAQKLGVEQPTLYWHVKNKRALLDALAIEML 60  
Db 1 MSRLDKSVNSALELLNEVGIEGLTRKLAQKLGVEQPTLYWHVKNKRALLDALAIEML 60  
QY 61 DRHHTFCPLEGESWQDFLRNKAQKSFRCALLSHRDGAKVHLGTRPTEKQYETLENOIAFL 120  
Db 61 DRHHTFCPLEGESWQDFLRNKAQKSFRCALLSHRDGAKVHLGTRPTEKQYETLENOIAFL 120  
QY 121 CQQGFSENLALYALSAYGHFTLGCVLDDQEHQVAKEREPTTDSMPPLLRQAIELFDHQ 180  
Db 121 CQQGFSENLALYALSAYGHFTLGCVLDDQEHQVAKEREPTTDSMPPLLRQAIELFDHQ 180  
QY 181 GAEPAPFLGLELIICGLEKQKCESGS 207  
Db 181 GAEPAPFLGLELIICGLEKQKCESGS 207

RESULT 2  
R89763 R89763 standard; Protein; 207 AA.  
AC R89763:  
DT 07-JUL-1996 (first entry)  
DE Wild type Tn-10 derived Tet repressor.  
KW Tet repressor; Herpes simplex virus; HSV; virion protein 16; VP16;  
KW fusion protein; gene expression; regulation; inhibition; activation;  
KW transcription.  
OS Transposon Tn-10.  
PN W09601313-A1.  
PD 18-JAN-1996.  
PF 29-JUN-1995; U08179.  
PR 01-JUL-1994; US-270637.

Shukla  
09/24/1347  
08/14/86814  
Seg. 158 174 19

15-JUL-1994; US-273876.  
 03-FEB-1995; US-383754.  
 07-JUN-1995; US-486814.  
 (BUJA/) BUJARD H.  
 (GOSS/) GOSSSEN M.  
 PI Bujard H, Gossen M;  
 DR WPI; 96-087666/09.  
 DR N-PSDB; T11351.  
 PT New tetracycline-regulated transcription modulators - comprising  
 PT fusion proteins which bind to tet operator sequences to activate or  
 PT inhibit transcription  
 PS Claim 38; Page 78; 112pp; English.  
 CC Fusion proteins comprising a first polypeptide which binds to a tet  
 CC operator sequence in the presence of tetracycline or a tetracycline  
 CC analogue, operatively linked to a second polypeptide which either  
 CC activates or inhibits transcription in eukaryotic cells may be used  
 CC to activate or inhibit transcription. Such proteins may be used to  
 CC regulate gene expression in cells and may be particularly useful for  
 CC gene therapy and for expression of gene products in transgenic  
 CC organisms. Induction of gene expression is rapid, efficient and  
 CC strong, typically 1000-2000 fold. The inducing agent does not cause  
 CC pleiotropic effects or cytotoxicity in eukaryotic cells. This  
 CC sequence is the wild type Tn-10 tet repressor.  
 SQ Sequence 207 AA;

Query Match 100.0%; Score 1080; DB 1; Length 207;  
 Best Local Similarity 100.0%; Pred No. 5.9e-113; Indels 0; Gaps 0;  
 Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSRLDKSVINSALLELNEVIGIEGLTRKLAQKLGVEQPTLYWHVKNKRALDLALAIEM 60  
 DB 1 MSRLDKSVINSALLELNEVIGIEGLTRKLAQKLGVEQPTLYWHVKNKRALDLALAIEM 60  
 QY 61 DRHHTFCPLEGESWQDFLNKAKSFRCALLSHRDGAKVHLGTRPTEKQYETLENQALFL 120  
 DB 61 DRHHTFCPLEGESWQDFLNKAKSFRCALLSHRDGAKVHLGTRPTEKQYETLENQALFL 120  
 QY 121 CQGFSLLENALYALS AVGHFTLCVLEDOHQVAKEREETPTTDSMPPLLRQAIELFDHQ 180  
 DB 121 CQGFSLLENALYALS AVGHFTLCVLEDOHQVAKEREETPTTDSMPPLLRQAIELFDHQ 180  
 QY 181 GAEPALFGLLELICGLKOLKCESGS 207  
 DB 181 GAEPALFGLLELICGLKOLKCESGS 207

RESULT 3  
 ID W08296 standard; Protein; 207 AA.  
 AC W08296;  
 DT 19-MAR-1997 (first entry)  
 DE Wild-type E.coli Tn10-derived Tet repressor.  
 KW Tet; tetracycline; Tc; operator; transcription; regulation; inducible;  
 KW repressor; gene expression; therapy; transgenic animal; disease model;  
 KW Tn10; transposon 10.  
 OS Escherichia coli.  
 PN W08640892-Al.  
 PD 19-DEC-1996.  
 PF 06-JUN-1996; U09049.  
 PR 07-JUN-1995; US-485971.  
 PA (BADI) BASF AG.  
 PI Bujard H, Gossen M, Helbl V, Hillen W, Schnappinger D;  
 DR WPI; 97-052305/05.  
 PT Nucleic acid encoding tetracycline-inducible transcription  
 PT regulatory fusion protein - comprising modified tetracycline  
 PT repressor able to bind mutant tet operator, fused to transcription  
 PT regulator, useful for modulating eukaryotic gene expression  
 PS Example 4; Page 84-85; 117pp; English.  
 CC W08296 represents wild-type tet repressor (TetR) protein, derived from  
 CC transposon 10 (Tn10) of E. coli. The wild-type sequence is used for  
 CC the production of a mutated TetR (e.g. so as to bind its target in the  
 CC presence rather than the absence of tetracycline). Mutant TetR proteins

CC can be fused to a transcriptional activator e.g. VP16 (herpes simplex  
 CC virus virion protein 16), to form a Tc-controlled transcriptional  
 CC activator (tTA). The main invention of the specification concerns  
 CC modified TetR proteins that bind to modified class B tet operator  
 CC sequences tetO-4C and tetO-6C (see T45711 and T48478). Modified TetR  
 CC proteins can be fused to any transcription regulatory polypeptide and  
 CC used to control transcription of a tetO-4C or tetO-6C linked gene.  
 CC Nucleic acid encoding such a fusion protein may be introduced into a  
 CC cell and transcription of the protein can be controlled by altering the  
 CC concn of tetracycline (or an analogue) in the cell, as appropriate.  
 CC This ability to modulate gene expression in a predictable way is very  
 CC useful in gene therapy and for recombinant protein prodn. in cultured  
 CC cells or transgenic animals. The Tc-inducible system is also useful for  
 CC the prodn. of transgenic animal models for the study of disease and  
 CC also for the study of gene function e.g. during differentiation. The  
 CC Tc-inducible system allows rapid activation of gene transcription  
 CC without cellular toxicity, high concns. of inducer are not required.  
 SQ Sequence 207 AA;

Query Match 100.0%; Score 1080; DB 1; Length 207;  
 Best Local Similarity 100.0%; Pred No. 5.9e-113; Indels 0; Gaps 0;  
 Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSRLDKSVINSALLELNEVIGIEGLTRKLAQKLGVEQPTLYWHVKNKRALDLALAIEM 60  
 DB 1 MSRLDKSVINSALLELNEVIGIEGLTRKLAQKLGVEQPTLYWHVKNKRALDLALAIEM 60  
 QY 61 DRHHTFCPLEGESWQDFLNKAKSFRCALLSHRDGAKVHLGTRPTEKQYETLENQALFL 120  
 DB 61 DRHHTFCPLEGESWQDFLNKAKSFRCALLSHRDGAKVHLGTRPTEKQYETLENQALFL 120  
 QY 121 CQGFSLLENALYALS AVGHFTLCVLEDOHQVAKEREETPTTDSMPPLLRQAIELFDHQ 180  
 DB 121 CQGFSLLENALYALS AVGHFTLCVLEDOHQVAKEREETPTTDSMPPLLRQAIELFDHQ 180  
 QY 181 GAEPALFGLLELICGLKOLKCESGS 207  
 DB 181 GAEPALFGLLELICGLKOLKCESGS 207

RESULT 4  
 ID W48357 standard; Protein; 651 AA.  
 AC W48357;  
 DT 20-JUL-1998 (first entry)  
 DE Multi-chimeric transactivating factor tTAER fusion protein.  
 KW Multi-chimeric transactivating factor; tTAER; tetR;  
 KW tetracycline repressor; HSV; oestrogen receptor; promoter;  
 KW packaging cell line; retrovirus; retroviral particle; vector;  
 KW gene delivery; gene therapy.  
 OS Chimeric - Escherichia coli.  
 OS Chimeric - Herpes simplex virus.  
 OS Chimeric - Mammalia.  
 PN W09805759-Al.  
 PD 12-FEB-1998.  
 PF 06-AUG-1997; U13846.  
 PR 07-AUG-1996; US-594652.  
 PA (CITY) CITY OF HOPE.  
 PA (REGC) UNIV CALIFORNIA.  
 PI Chen S, Friedmann T, Miyanochara A, Prussak CE, Yee J;  
 DR WPI; 98-145602/13.  
 DR N-PSDB; V17756.  
 PT New packaging cell lines for pseudotyped retroviral vectors -  
 PT comprises sequences encoding retroviral gag and pol polypeptide(s)  
 PT and envelope protein  
 PS Disclosure; Fig 4A-C; 84pp; English.  
 CC This fusion protein comprises a multi-chimeric transactivator,  
 CC designated tTAER, that is composed of (from the N-terminus to the  
 CC C-terminus) the Escherichia coli tetR polypeptide, the  
 CC transcriptional activation domain of herpes simplex virus VP16, and  
 CC the ligand binding domain of the oestrogen receptor. An inducible  
 CC expression system of the invention is composed of tTAER and a

Fri Jun 11 11:58:05 1999

CC minimal promoter (see V17755) derived from the immediate early gene  
 CC of cytomegalovirus linked to 7 tandem copies of the tet operator  
 CC (teto) that is the binding site for tetr, which in turn can be  
 CC linked to a nucleotide sequence of interest. The invention relates  
 CC to packaging cell lines derived from HeLa, D17, MDCK, BHK or  
 CC preferably Cf2Th cells and recombinant retroviral particles  
 CC produced by them, particularly pseudotyped retroviral particles.  
 CC Retroviral particles are produced by inducibly expressing an  
 CC envelope protein by linking an envelope protein-encoding nucleotide  
 CC sequence to the inducible expression system. The products can be  
 CC used for the inducible expression in cells of polypeptides, e.g.  
 CC cytotoxic products or therapeutic agents. The activation of the  
 CC inducible expression system requires 2 independent signals, thus  
 CC reducing the incidence of undesired transcriptional activation.  
 SQ Sequence 651 AA;

Query Match 100.0%; Score 1080; DB 1; Length 651;  
 Best Local Similarity 100.0%; Pred. No. 2.9e-112;  
 Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSRLDKSVINSALELNEVGIEGLTRKLAQKLGVEQPTLYWHVKNKRALDLAIEML 60  
 DB 1 MSRLDKSVINSALELNEVGIEGLTRKLAQKLGVEQPTLYWHVKNKRALDLAIEML 60

QY 61 DRHTHFCPLGEGSWQDFLNKAKSFRCALLSHRDGAKVHLGTRPTKEQYETLENQALFL 120  
 DB 61 DRHTHFCPLGEGSWQDFLNKAKSFRCALLSHRDGAKVHLGTRPTKEQYETLENQALFL 120

QY 121 CQGFSLLENALYALS AVGHFTLGCVLDEQHQVAKEREPTTDSMPPLLRQAIELFDHQ 180  
 DB 121 CQGFSLLENALYALS AVGHFTLGCVLDEQHQVAKEREPTTDSMPPLLRQAIELFDHQ 180

QY 181 GAEPFLFGLLEIICGLEKQKCESGS 207  
 DB 181 GAEPFLFGLLEIICGLEKQKCESGS 207

RESULT 5  
 ID W47583 standard; Protein; 651 AA.  
 AC W47583; 1998 (first entry)  
 DT Multi-chimeric transactivator tTAER sequence.  
 DE Inducible expression system; modulation; cytotoxin; therapeutic;  
 KW tTAER; multi-chimeric; transactivator; tetR; Vp16;  
 KW ligand-binding domain.  
 OS Escherichia coli.  
 OS Herpes simplex virus.  
 PN WO9805754-A2.  
 PD 12-FEB-1998.  
 PF 29-JUL-1997; U13221.  
 PR 07-AUG-1996; US-693940.  
 PA (CITY) CITY OF HOPE.  
 PA (REGC) UNIV CALIFORNIA.  
 PI Chen ST, Friedmann T, Yee JK;  
 DR WPI; 98-145597/13.  
 DR N-PSDB; V18690.  
 PT New inducible expression systems - comprising multi-chimeric  
 PT transactivator, induces transcription from promoter in the absence  
 PT of first ligand and presence of second ligand  
 PS Example 1; Fig 4; 72pp; English.  
 CC The sequence is that of multi-chimeric transactivator tTAER.  
 CC This is a fusion protein comprising, from N-terminus to C-terminus,  
 CC the E. coli tetr polypeptide, the transcription activation domain  
 CC of HSV Vp16 and the ligand-binding domain of an oestrogen receptor  
 CC from pHE14 (Kumar et al. 1986). DNA encoding it can be used as part of  
 CC an inducible expression system which induces transcription from the  
 CC promoter in the absence of the first ligand and presence of the  
 CC second. The products can be used for the inducible expression in cells  
 CC of polypeptides such as cytotoxic products or therapeutic products.  
 CC The activation of the inducible expression system requires 2  
 CC independent signals, which reduces the incidence of undesired

CC transcriptional activation.  
 SQ Sequence 651 AA;

Query Match 100.0%; Score 1080; DB 1; Length 651;  
 Best Local Similarity 100.0%; Pred. No. 2.9e-112;  
 Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSRLDKSVINSALELNEVGIEGLTRKLAQKLGVEQPTLYWHVKNKRALDLAIEML 60  
 DB 1 MSRLDKSVINSALELNEVGIEGLTRKLAQKLGVEQPTLYWHVKNKRALDLAIEML 60

QY 61 DRHTHFCPLGEGSWQDFLNKAKSFRCALLSHRDGAKVHLGTRPTKEQYETLENQALFL 120  
 DB 61 DRHTHFCPLGEGSWQDFLNKAKSFRCALLSHRDGAKVHLGTRPTKEQYETLENQALFL 120

QY 121 CQGFSLLENALYALS AVGHFTLGCVLDEQHQVAKEREPTTDSMPPLLRQAIELFDHQ 180  
 DB 121 CQGFSLLENALYALS AVGHFTLGCVLDEQHQVAKEREPTTDSMPPLLRQAIELFDHQ 180

QY 181 GAEPFLFGLLEIICGLEKQKCESGS 207  
 DB 181 GAEPFLFGLLEIICGLEKQKCESGS 207

RESULT 6  
 ID W71326 standard; Protein; 207 AA.  
 AC W71326;  
 DT 04-DEC-1998 (first entry)  
 DE Protein encoded by wild type Tn10-derived tet repressor.  
 KW Tet repressor; tetracycline; regulation; expression;  
 KW Tet operator-linked gene; tet operator.  
 OS Unidentified.  
 PN US814618-A.  
 PD 29-SEP-1998.  
 PF 07-JUN-1995; 485978.  
 PR 07-JUN-1995; US-485978.  
 PR 14-JUN-1993; US-076327.  
 PR 14-JUN-1993; US-076726.  
 PR 14-JUN-1994; US-260452.  
 PR 01-JUL-1994; US-270637.  
 PR 15-JUL-1994; US-275876.  
 PR 06-FEB-1995; US-383754.  
 PA (BADI) BASF AG.  
 PA (KNOL) KNOLL AG.  
 PI Bujard H, Gossen M;  
 DR WPI; 98-541795/46.  
 DR N-PSDB; V60088.  
 PT Tetracycline based regulation of gene expression - uses a  
 PT tetracycline operator sequence joined to a gene of interest, the  
 PT gene of interest being induced in the presence, but not absence of  
 PT the antibiotic  
 PS Claim 4; Columns 77-80; 63pp; English.  
 CC The present sequence is encoded by wild type Tn10-derived tet repressor.  
 CC It is used in the course of the invention. The specification describes a  
 CC method for regulating expression of a Tet (tetracycline) operator-linked  
 CC gene in a cell of a subject. The method comprises introducing into the  
 CC cell a nucleic acid encoding a fusion protein which inhibits  
 CC transcription in eukaryotic cells, the fusion protein comprising a  
 CC polypeptide which binds to a Tet operator sequence, operatively linked  
 CC to heterologous second polypeptide which inhibits transcription in  
 CC eukaryotic cells and modulating the concentration of a tetracycline  
 CC (analogue) in the subject. The method is used for the regulation of  
 CC gene expression system, using tetracycline (analogues). The system  
 CC enables a gene coupled to the system to be induced in the presence of  
 CC Tet and then stopped when Tet is removed.  
 SQ Sequence 207 AA;

Query Match 100.0%; Score 1080; DB 1; Length 207;  
 Best Local Similarity 100.0%; Pred. No. 5.9e-113;  
 Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSRLDKSVINSALELLNEVGIEGLTRKLAQKLGVEQPTLYWHVKNKRALLDALATEML 60  
 DB 1 MSRLDKSVINSALELLNEVGIEGLTRKLAQKLGVEQPTLYWHVKNKRALLDALATEML 60  
 QY 61 DRHHTFCPLGESWQDFLRNKAQKSFRCALLSHRDGAKVHLGTRPTEKQYETLENQLAFL 120  
 DB 61 DRHHTFCPLGESWQDFLRNKAQKSFRCALLSHRDGAKVHLGTRPTEKQYETLENQLAFL 120  
 QY 121 COQGSLENALYALS AVGHFTLGCVLDEQHOVAKEERETPTTDSMPPLLRQAIELEFDHQ 180  
 DB 121 COQGSLENALYALS AVGHFTLGCVLDEQHOVAKEERETPTTDSMPPLLRQAIELEFDHQ 180  
 QY 181 GAEPALFGLLELIICGLEKQKCESGS 207  
 DB 181 GAEPALFGLLELIICGLEKQKCESGS 207

RESULT 7  
 R48630  
 ID R48630 standard; Protein; 207 AA.  
 AC R48630;  
 DT 19-AUG-1994 (first entry)  
 DE Sequence of the tetracycline repressor tetr.  
 KW Tetracycline repressor; TetR; Tn10; transposon 10.  
 OS Escherichia coli.  
 PN WO9404672-A.  
 PD 03-MAR-1994.  
 PF 26-AUG-1993; U08230.  
 PR 26-AUG-1992; US-935763.  
 PA (DNXD-) DNK CORP.  
 PI Byrne G;  
 DR WPI; 94-083191/10.  
 DR N-PSDB; Q56710.  
 PT Tetracycline repressor-mediated regulation system - useful for  
 PT controlling gene expression in transgenic animals  
 PS Disclosure; Page 41-42; 76pp; English.  
 CC The inventors claim a construct which comprises an animal promoter  
 CC element having a tetracycline repressor (tetR) operator. The  
 CC promoter element may be the PEPC promoter which is tissue specific  
 CC being expressed selectively in the liver and becoming active  
 CC shortly prior to birth. The tetR sequence in the  
 CC construct is 3' to a TATA-box sequence and is inserted into the NheI  
 CC site of the PEPC promoter element. The entire sequence of the tetR  
 CC repressor is given in Q56710/R48630.  
 SQ Sequence 207 AA;

Query Match 99.5%; Score 1075; DB 1; Length 207;  
 Best Local Similarity 99.5%; Pred. No. 2.1e-112;  
 Matches 206; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSRLDKSVINSALELLNEVGIEGLTRKLAQKLGVEQPTLYWHVKNKRALLDALATEML 60  
 DB 1 MSRLDKSVINSALELLNEVGIEGLTRKLAQKLGVEQPTLYWHVKNKRALLDALATEML 60  
 QY 61 DRHHTFCPLGESWQDFLRNKAQKSFRCALLSHRDGAKVHLGTRPTEKQYETLENQLAFL 120  
 DB 61 DRHHTFCPLGESWQDFLRNKAQKSFRCALLSHRDGAKVHLGTRPTEKQYETLENQLAFL 120  
 QY 121 COQGSLENALYALS AVGHFTLGCVLDEQHOVAKEERETPTTDSMPPLLRQAIELEFDHQ 180  
 DB 121 COQGSLENALYALS AVGHFTLGCVLDEQHOVAKEERETPTTDSMPPLLRQAIELEFDHQ 180  
 QY 181 GAEPALFGLLELIICGLEKQKCESGS 207  
 DB 181 GAEPALFGLLELIICGLEKQKCESGS 207

RESULT 8  
 R64809  
 ID R64809 standard; Protein; 297 AA.  
 AC R64809;

DT 17-JUL-1995 (first entry)  
 DE tTAs transactivator.  
 KW tTAs; transactivator; tetracycline-controllable transactivator;  
 KW conditional inactivation; homologous recombination; gene expression;  
 KW gene regulation; gene therapy; tetracycline-resistance; tetR;  
 KW transgenic animal.  
 OS Herpes simplex virus K12, KOS.  
 PN WO9429442-A.  
 PD 22-DEC-1994.  
 PF 14-JUN-1994; U06734.  
 PR 14-JUN-1993; US-076327.  
 PA (BADI) BASF AG.  
 PI Bujard H, Gossen M, Salfeld JG, Voss JW;  
 DR WPI; 95-036472/05.  
 DR N-PSDB; Q76265.  
 PT Regulatory systems using tetracycline-controllable transactivator  
 PT (tTA) - useful for conditional inactivation or modulation of  
 PT gene expression in a host cell or animal  
 PS Disclosure; Page 50-51; 103pp; English.  
 CC A DNA fragment of pMSVP16 coding for the C-terminal 97 amino  
 CC acids of VP16 of HSV was blunted and inserted in pUHD14-1. The  
 CC resulting plasmid, pUHD151-1, encoded a tetR-VP16 fusion protein,  
 CC or tetracycline-controllable transactivator (smaller version, tTAs),  
 CC whose sequence is given in Q76265.  
 SQ Sequence 297 AA;

Query Match 99.5%; Score 1075; DB 1; Length 297;  
 Best Local Similarity 99.5%; Pred. No. 3.5e-112;  
 Matches 206; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSRLDKSVINSALELLNEVGIEGLTRKLAQKLGVEQPTLYWHVKNKRALLDALATEML 60  
 DB 1 MSRLDKSVINSALELLNEVGIEGLTRKLAQKLGVEQPTLYWHVKNKRALLDALATEML 60  
 QY 61 DRHHTFCPLGESWQDFLRNKAQKSFRCALLSHRDGAKVHLGTRPTEKQYETLENQLAFL 120  
 DB 61 DRHHTFCPLGESWQDFLRNKAQKSFRCALLSHRDGAKVHLGTRPTEKQYETLENQLAFL 120  
 QY 121 COQGSLENALYALS AVGHFTLGCVLDEQHOVAKEERETPTTDSMPPLLRQAIELEFDHQ 180  
 DB 121 COQGSLENALYALS AVGHFTLGCVLDEQHOVAKEERETPTTDSMPPLLRQAIELEFDHQ 180  
 QY 181 GAEPALFGLLELIICGLEKQKCESGS 207  
 DB 181 GAEPALFGLLELIICGLEKQKCESGS 207

RESULT 9  
 R85324  
 ID R85324 standard; Protein; 297 AA.  
 AC R85324;  
 DT 13-MAR-1996 (first entry)  
 DE tTAs transactivator.  
 KW Transactivator; tTAs; tet repressor; tetR; virion protein 16; VP16;  
 KW gene expression; tetracycline-responsive promoter;  
 KW transgenic animal.  
 OS Chimeric Escherichia coli;  
 OS Chimeric Herpes simplex virus.  
 PN US5464758-A.  
 PD 07-NOV-1995.  
 PR 14-JUN-1993; 076726.  
 PR 14-JUN-1993; US-076726.  
 PA (BUJA/) BUJARD H.  
 PA (GOSS/) GOSSEN M.  
 PI Bujard H, Gossen M;  
 DR WPI; 95-392612/50.  
 DR N-PSDB; T06868.  
 PT Polynucleotide encoding transactivator fusion protein contg. tet  
 PT repressor - used to control expression of gene regulated by minimal  
 PT promoter linked to tet operon, and vectors and cells where gene  
 PT expression is regulated by tetracycline  
 PS Disclosure; Fig 5A-B; 37pp; English.

CC A fusion protein (R85324) composed of the *Escherichia coli* trn10  
 CC tet repressor (tetR) and a 127-amino acid C-terminal portion of the  
 CC transcription activating domain of herpes simplex virus virion  
 CC protein 16 (VP16) is the product of novel tetracycline-controlled  
 CC transactivator tTA (T06868). The transactivator is used to regulate  
 CC expression of a heterologous gene operably linked to a minimal promoter  
 CC derived from human cytomegalovirus promoter IE (see T06869-70) and at  
 CC least 1 tet operator (tetO) sequence. On/off regulation of expression  
 CC of the heterologous protein by host eukaryotic cells is provided by  
 CC varying the medium tetracycline conc. Transgenic animals producing  
 CC a desired protein in their milk can also be produced.  
 CC Sequence 297 AA;  
 SQ

Query Match 99.5%; Score 1075; DB 1; Length 297;  
 Best Local Similarity 99.5%; Pred. NO. 3.5e-112;  
 Matches 206; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 MSRLDKSKVINSALLELLNEVGIEGLTTRKLAQKLGVEQPTLYWHVKNKRALDLAIAEML 60  
 DB 1 MSRLDKSKVINSALLELLNEVGIEGLTTRKLAQKLGVEQPTLYWHVKNKRALDLAIAEML 60  
 QY 61 DRHHTFCPLEGESWQDFLRNKAQKFRCALLSHRDGAKVHLGTRPTEKQYETLENQALFL 120  
 DB 61 DRHHTFCPLEGESWQDFLRNKAQKFRCALLSHRDGAKVHLGTRPTEKQYETLENQALFL 120  
 QY 121 CQOGFSLENALYALSANGVHFTLCVLEDOHQVAKEREETPTTDSMPPLLRQAIEFLDHQ 180  
 DB 121 CQOGFSLENALYALSANGVHFTLCVLEDOHQVAKEREETPTTDSMPPLLRQAIEFLDHQ 180  
 QY 181 GAEPALFGLLEIICGLEKQKCESGS 207  
 DB 181 GAEPALFGLLEIICGLEKQKCESGS 207

RESULT 10  
 R88636  
 ID R88636 standard; Protein; 1088 AA.  
 AC R88636;  
 DT 27-JUN-1996 (first entry)  
 DE Plasmid pASK75 open reading frame (c) translation.  
 KW Tetracycline; resistance; TetR gene; repressor protein; TetA gene;  
 KW prokaryotic expression vector; beta-lactamase promoter; bla gene;  
 KW circular; cyclic.  
 OS Synthetic.  
 FH Key Location/Qualifiers  
 FT misc\_difference 1..1088 /note="each X corresponds to a termination codon"  
 FT peptide 49..69 /label="OmpA\_leader"  
 FT protein 614..820 /label="Tetr"  
 FT region 601..609 /note="tetracycline repressor"  
 FT /note="C-terminal end of beta-lactamase; the  
 FT exact position of the N-terminus is not  
 FT given in the specification"  
 PI WO9532295-AL.  
 PN 30-NOV-1995.  
 PE 17-MAY-1995; E01862.  
 PR 19-MAY-1994; DE-417598.  
 PA (PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.  
 PI Skerra A, Wardenberg C;  
 DR WPI; 96-020588/02.  
 DR N-PSDB; T11192.  
 PT Prokaryotic vector for regulated prodn. of heterologous protein gene  
 PT - controlled by promoter repressed by tetracycline repressor  
 PT protein, prodn. of which is controlled by independent promoter not  
 PT regulated by this repressor  
 PS Disclosure; Fig 1a; 50pp; German.  
 CC Plasmid pASK75 is a specifically claimed example of a prokaryotic  
 CC expression vector comprising the tetracycline promoter/operator (P/O)  
 CC region and the tetracycline repressor (tetR) gene. The tetR gene is

CC under the control of the beta-lactamase (bla) promoter. Expression of  
 CC heterologous genes inserted downstream of the tet P/O is controlled  
 CC by the activity of the Tet repressor. Plasmid pASK75 was derived  
 CC from pASK60-Strep (Schmidt and Skerra, Protein. Eng. 6 (1993), 109-  
 CC 122) and further contains the OmpA leader peptide coding sequence, a  
 CC polylinker and a streptavidin-binding peptide strep-tag coding  
 CC region from the parent plasmid. In addition, pASK75 contains the  
 CC ColEI origin of replication, the bla gene, intergene regions of  
 CC filamentous phage f1 and a lipoprotein terminator.  
 CC The specification includes the translated sequences from all three  
 CC reading frames of pASK75; the OmpA leader, TetR and Bla coding  
 CC regions are all in reading frame (c).  
 SQ Sequence 1088 AA;

Query Match 99.5%; Score 1075; DB 1; Length 1088;  
 Best Local Similarity 99.5%; Pred. NO. 2.2e-111;  
 Matches 206; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 MSRLDKSKVINSALLELLNEVGIEGLTTRKLAQKLGVEQPTLYWHVKNKRALDLAIAEML 60  
 DB 614 MSRLDKSKVINSALLELLNEVGIEGLTTRKLAQKLGVEQPTLYWHVKNKRALDLAIAEML 673  
 QY 61 DRHHTFCPLEGESWQDFLRNKAQKFRCALLSHRDGAKVHLGTRPTEKQYETLENQALFL 120  
 DB 674 DRHHTFCPLEGESWQDFLRNKAQKFRCALLSHRDGAKVHLGTRPTEKQYETLENQALFL 733  
 QY 121 CQOGFSLENALYALSANGVHFTLCVLEDOHQVAKEREETPTTDSMPPLLRQAIEFLDHQ 180  
 DB 734 CQOGFSLENALYALSANGVHFTLCVLEDOHQVAKEREETPTTDSMPPLLRQAIEFLDHQ 793  
 QY 181 GAEPALFGLLEIICGLEKQKCESGS 207  
 DB 794 GAEPALFGLLEIICGLEKQKCESGS 820

RESULT 11  
 W21994  
 ID W21994 standard; Protein; 354 AA.  
 AC W21994;  
 DT 26-JAN-1998 (first entry)  
 DE Tetracycline transactivator fusion protein tetRIEIA.  
 KW Baculovirus; recombinant protein; neurotoxin; parasite;  
 KW biological control; insecticide; insecticidal protein;  
 KW nuclear polyhedrosis virus; early promoter; tetracycline.  
 OS Chimeric - Baculovirus.  
 OS Chimeric - *Escherichia coli*.  
 FH Key Location/Qualifiers  
 FT Peptide 1..209 /label="tetr peptide"  
 FT Peptide 210..354 /label="IEIA peptide"  
 FT WO9723636-AL.  
 PN 03-JUL-1997.  
 PE 16-DEC-1996; U20512.  
 PR 12-DEC-1996; US-764369.  
 PR 22-DEC-1995; US-009120.  
 PA (DUPO) DU PONT DE NEMOURS & CO E. I.  
 PI McCutchen BF;  
 DR WPI; 97-351071/32.  
 DR Controlling expression of insecticidal proteins - using recombinant  
 PT baculovirus containing tetracycline transactivator protein  
 PS Example 2; Fig 8; 63pp; English.  
 CC This transactivator protein is encoded by a hybrid gene consisting of  
 CC the tetracycline repressor gene (tetR) of *E. coli* (T74611) fused  
 CC in frame with the immediate early promoter (IEIA) of baculovirus  
 CC (T74614). The protein is used to demonstrate a novel method of gene  
 CC expression regulation of recombinant baculoviruses in an insect  
 CC cell/host. An insect cell is constructed having a chimeric gene, which  
 CC consists of a promoter (e.g. T74614) operably linked to DNA encoding a  
 CC regulatory protein capable of affecting gene expression. The regulatory  
 CC protein is preferably a tetracycline transactivator protein  
 CC (e.g. W21994). A recombinant baculovirus expression vector is then

CC constructed. This has a second chimeric gene, which consists of a  
 CC promoter (one or more tetracycline operator sites operably linked to a  
 CC minimal promoter region, see T74615-16) that is affected by the above  
 CC regulatory protein. This promoter is operably linked to DNA encoding an  
 CC insecticidal protein, (w21993). The recombinant baculovirus expression  
 CC vector is introduced into the recombinant insect cell. Expression of the  
 CC regulatory protein affects expression of the insecticidal protein. The  
 CC recombinant insect cells and the recombinant baculoviruses may be used  
 CC in the production of insecticidal recombinant baculoviruses. Insect cells  
 CC treated with recombinant viruses, as described, show an increase in the  
 CC yield of polyhedral inclusion bodies.  
 SQ Sequence 354 AA;

Query Match 99.5%; Score 1075; DB 1; Length 354;  
 Best Local Similarity 99.5%; Pred. No. 4.5e-112;  
 Matches 206; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 MSRLDKSVINSALLELLNEVGIEGLTTRKLAQKLGVEOPTLYWHVKNKRALLDALAIEML 60  
 DB 1 MSRLDKSVINSALLELLNEVGIEGLTTRKLAQKLGVEOPTLYWHVKNKRALLDALAIEML 60  
 QY 61 DRHHTHFCPLGESQWDFLRNKAISFRCALLSHRDGAKVHLGTRPTKEQYETLENQALFL 120  
 DB 61 DRHHTHFCPLGESQWDFLRNKAISFRCALLSHRDGAKVHLGTRPTKEQYETLENQALFL 120  
 QY 121 COQGSLENALYALSAGVGHFTLGCVLDEQEHQVAKEREETPTTDSMPPLLRQAIELEFDHQ 180  
 DB 121 COQGSLENALYALSAGVGHFTLGCVLDEQEHQVAKEREETPTTDSMPPLLRQAIELEFDHQ 180  
 QY 181 GAEPAPFLGLELIICGLEKQKLCESGS 207  
 DB 181 GAEPAPFLGLELIICGLEKQKLCESGS 207

RESULT 12  
 W54311  
 ID W54311 standard; Protein; 349 AA.  
 AC W54311;  
 DE 28-AUG-1998 (first entry)  
 DT Pleckstrin homology domain-tet repressor fusion.  
 KW Pleckstrin homology domain; PH domain; peptide library;  
 KW functional molecular surface; protein structural template;  
 KW vaccine; gene therapy; cytohesin 1; human; plasmid pTL2;  
 KW tet repressor.  
 OS Chimeric - Homo sapiens.  
 OS Chimeric - Escherichia coli.  
 OS Synthetic.  
 FH Key Location/Qualifiers  
 FT Protein 1..221  
 FT /label= Tet\_repressor  
 FT 222..349  
 FT /label= PH\_domain  
 PN W09745538-A1.  
 PD 04-DEC-1997.  
 PF 30-MAY-1997; E02840.  
 PR 31-MAY-1996; EP-108776.  
 PA (MEDI-) MEDIGENE AG.  
 PI Bruhn H, Funk M, Henkel T, Steipe B;  
 DR WPI: 98-230215/20.  
 DR N-PSDB: V26524.  
 PT Vectors used to produce PH domain-like peptide libraries - which are  
 PT screened for therapeutically useful peptide(s), e.g. to produce  
 PT vaccines  
 PS Example 4; Page 79-81; 137pp; English.  
 CC This polypeptide comprises a fusion between a synthetic Pleckstrin  
 CC homology (PH) domain (see W54285) of human cytohesin 1 and a tet  
 CC repressor. It is encoded by plasmid pTL2 (see V26524). The fusion  
 CC protein is expressed from a polycistronic message, constitutively  
 CC transcribed from the beta-lactamase promoter. The plasmid was used  
 CC to transform Escherichia coli JM109 cells. The construct was used  
 CC to demonstrate the folding, stability, functionality and ease of  
 CC handling of a fusion protein comprising a synthetic PH domain and

CC a DNA binding domain. The invention provides vectors that are used  
 CC for the production of PH domain-like peptide libraries, which can  
 CC be screened to identify peptides that have desirable properties,  
 CC especially novel binding or catalytic properties, and which may be  
 CC of use in research or therapy, or as vaccines. Novel synthetic  
 CC protein structural templates for the generation, screening and  
 CC evolution of functional molecular surfaces are provided.  
 SQ Sequence 349 AA;

Query Match 99.5%; Score 1075; DB 1; Length 349;  
 Best Local Similarity 99.5%; Pred. No. 4.4e-112;  
 Matches 206; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 MSRLDKSVINSALLELLNEVGIEGLTTRKLAQKLGVEOPTLYWHVKNKRALLDALAIEML 60  
 DB 2 MSRLDKSVINSALLELLNEVGIEGLTTRKLAQKLGVEOPTLYWHVKNKRALLDALAIEML 61  
 QY 61 DRHHTHFCPLGESQWDFLRNKAISFRCALLSHRDGAKVHLGTRPTKEQYETLENQALFL 120  
 DB 62 DRHHTHFCPLGESQWDFLRNKAISFRCALLSHRDGAKVHLGTRPTKEQYETLENQALFL 121  
 QY 121 COQGSLENALYALSAGVGHFTLGCVLDEQEHQVAKEREETPTTDSMPPLLRQAIELEFDHQ 180  
 DB 122 COQGSLENALYALSAGVGHFTLGCVLDEQEHQVAKEREETPTTDSMPPLLRQAIELEFDHQ 181  
 QY 181 GAEPAPFLGLELIICGLEKQKLCESGS 207  
 DB 182 GAEPAPFLGLELIICGLEKQKLCESGS 208

RESULT 13  
 R85323  
 ID R85323 standard; Protein; 335 AA.  
 AC R85323;  
 DT 13-MAR-1996 (first entry)  
 DE TTA transactivator.  
 KW Transactivator; tTA; tet repressor; tetR; virion protein 16; VP16;  
 KW gene expression; tetracycline-responsive promoter;  
 KW transgenic animal.  
 OS Chimeric Escherichia coli;  
 OS Chimeric Herpes simplex virus.  
 PN US5464758-A.  
 PD 07-NOV-1995.  
 PF 14-JUN-1993; 076726.  
 PR 14-JUN-1993; US-076726.  
 PA (BUJA) BUJARD H.  
 PA (GOSSE) GOSSEN M.  
 PI Bujard H, Gossen M;  
 DR WPI: 95-392612/50.  
 DR N-PSDB: T06867.  
 PT Polynucleotide encoding trans:activator fusion protein contg. tet  
 PT repressor - used to control expression of gene regulated by minimal  
 PT promoter linked to tet operon, and vectors and cells where gene  
 PT expression is regulated by tetracycline  
 PS Disclosure; Fig 4A-B; 37pp; English.  
 CC A fusion protein (R85323) composed of the Escherichia coli Tn10  
 CC tet repressor (tetR) and a 127-amino acid C-terminal portion of the  
 CC transcription activating domain of herpes simplex virus virion  
 CC protein 16 (VP16) is the product of novel tetracycline-controlled  
 CC transactivator tTA (T06867). The transactivator is used to regulate  
 CC expression of a heterologous gene operably linked to a minimal promoter  
 CC derived from human cytomegalovirus promoter IE (see T06869-70) and at  
 CC least 1 tet operator (tetO) sequence. On/off regulation of expression  
 CC of the heterologous protein by host eukaryotic cells is provided by  
 CC varying the medium tetracycline conc. Transgenic animals producing  
 CC a desired protein in their milk can also be produced.  
 SQ Sequence 335 AA;

Query Match 99.4%; Score 1073; DB 1; Length 335;  
 Best Local Similarity 99.5%; Pred. No. 7e-112;  
 Matches 206; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

W08323 standard; Protein; 207 AA.  
 AC W08323:  
 DT 21-MAR-1997 (first entry)  
 DE Mutated tet repressor for Tet-inducible transcriptional regulator.  
 KW Tet; tetracycline; rc; operator; transcription; regulation; inducible;  
 KW repressor; gene expression; therapy; transgenic animal; disease model.  
 OS Synthetic.  
 FH Key Location/Qualifiers  
 FT misc\_difference 37 /label= substitution  
 FT /note= "Lys replaces wild-type Glu"  
 FT misc\_difference 39 /label= substitution  
 FT /note= "Gln replaces wild-type Pro"  
 PN W09640892-Al.  
 PD 19-DEC-1996.  
 PF 06-JUN-1996; U09049.  
 PR 07-JUN-1995; US-485971.  
 PA (BADI) BASF AG.  
 PI Bujard H, Gossen M, Helbl V, Hillen W, Schnappinger D;  
 DR WPI; 97-052305/05.  
 PT Nucleic acid encoding tetracycline-inducible transcription  
 PT regulatory fusion protein - comprising modified tetracycline  
 PT repressor able to bind mutant tet operator, fused to transcription  
 PT regulator, useful for modulating eukaryotic gene expression  
 PT Claim 5; Page -; 117pp; English.  
 PS W08300-W08326 are mutated Tet repressor (TetR) proteins. They were  
 CC used to create fusion protein, comprising a mutated tetR fused to a  
 CC transcriptional control sequence (e.g. v-erbA silencer domain), to  
 CC produce tetracycline (Tc)-controlled transcriptional activators (ETA).  
 CC The main invention of the specification concerns modified TetR proteins  
 CC that bind to modified class B tet operator sequences tetO-4C and  
 CC tetO-6C (see T45711 and T48478). Modified TetR proteins can be fused  
 CC to any transcription regulatory polypeptide and used to control  
 CC transcription of a tetO-4C or tetO-6C linked gene. Nucleic acid  
 CC encoding such a fusion protein may be introduced into a cell and  
 CC transcription of the protein can be controlled by altering the concn.  
 CC of tetracycline (or an analogue) in the cell, as appropriate. This  
 CC ability to modulate gene expression in a predictable way is very useful  
 CC in gene therapy and for recombinant protein prodn. in cultured cells or  
 CC transgenic animals. The Tc-inducible system is also useful for the  
 CC prodn. of transgenic animal models for the study of disease and also  
 CC for the study of gene function e.g. during differentiation. The  
 CC Tc-inducible system allows rapid activation of gene transcription  
 CC without cellular toxicity, high concns. of inducer are not required.  
 SQ Sequence 207 AA;

RESULT 14

W08474  
 AC W08474 standard; Protein; 336 AA.  
 DT 28-MAR-1997 (first entry)  
 DE pTet-trak encoded protein.  
 KW pTet-trak; pTet-Splice; tetracycline transactivator; promoter;  
 KW gene expression; transgenic animal; animal model; drug screening.  
 OS Synthetic.  
 PN W09640946-Al.  
 PD 19-DEC-1996.  
 PF 07-JUN-1996; U10109.  
 PR 07-JUN-1995; US-474169.  
 PA (UYA) UNIV YALE.  
 PI Schatz DG;  
 DR WPI; 97-077273/07.  
 DR N-PSDB; T49877.  
 PT Nucleic acid encoding tetracycline transactivator fusion protein -  
 PT provides rapid and reversible control of gene expression, e.g. for  
 PT creating animal models for drug screening  
 PT Example 1; Fig 10b-g; 82pp; English.  
 PS Autoregulatory plasmid pTet-trak (T49877) contains a modified  
 CC tetracycline transactivator gene, trak, under control of a minimal  
 CC promoter. It encodes a 366-amino acid protein (W08474). An  
 CC autoregulatory tetracycline-regulated system for inducible gene  
 CC expression in eukaryotes has been developed that allows the rapid,  
 CC reversible control of gene expression in eukaryotic cells, and  
 CC creation of transgenic animals in which expression of a luciferase  
 CC reporter gene is controlled by altering the concentration of  
 CC tetracycline in the drinking water of the animals.  
 SQ Sequence 336 AA;

Query Match 99.1%; Score 1070; DB 1; Length 336;  
 Best Local Similarity 99.5%; Pred. No. 1.5e-111;  
 Matches 205; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 SRLDQSKVINSALLELLNEVGIEGLTRKLAQKLGVEOPTLYWHVKNKRALDALAEMLD 61  
 DB 3 SRLDQSKVINSALLELLNEVGIEGLTRKLAQKLGVEOPTLYWHVKNKRALDALAEMLD 62  
 QY 62 RHHTHFCPLGESWQDFLRNKAQKSFRCALLSHRDGAKVHLGTRPTEKQYETLENQALFL 121  
 DB 63 RHHTHFCPLGESWQDFLRNKAQKSFRCALLSHRDGAKVHLGTRPTEKQYETLENQALFL 122  
 QY 122 CQGFSLNALLYALSAGVHFTLCGVLEDOHQVAKEREPTTDSMPPLRQALIEFDHQ 181  
 DB 123 CQGFSLNALLYALSAGVHFTLCGVLEDOHQVAKEREPTTDSMPPLRQALIEFDHQ 182  
 QY 182 GAEPALFGLLEIICGLEKQKCESGS 207  
 DB 183 GAEPALFGLLEIICGLEKQKCESGS 208

RESULT 15

W08323 standard; Protein; 207 AA.  
 AC W08323:  
 DT 21-MAR-1997 (first entry)  
 DE Mutated tet repressor for Tet-inducible transcriptional regulator.  
 KW Tet; tetracycline; rc; operator; transcription; regulation; inducible;  
 KW repressor; gene expression; therapy; transgenic animal; disease model.  
 OS Synthetic.  
 FH Key Location/Qualifiers  
 FT misc\_difference 37 /label= substitution  
 FT /note= "Lys replaces wild-type Glu"  
 FT misc\_difference 39 /label= substitution  
 FT /note= "Gln replaces wild-type Pro"  
 PN W09640892-Al.  
 PD 19-DEC-1996.  
 PF 06-JUN-1996; U09049.  
 PR 07-JUN-1995; US-485971.  
 PA (BADI) BASF AG.  
 PI Bujard H, Gossen M, Helbl V, Hillen W, Schnappinger D;  
 DR WPI; 97-052305/05.  
 PT Nucleic acid encoding tetracycline-inducible transcription  
 PT regulatory fusion protein - comprising modified tetracycline  
 PT repressor able to bind mutant tet operator, fused to transcription  
 PT regulator, useful for modulating eukaryotic gene expression  
 PT Claim 5; Page -; 117pp; English.  
 PS W08300-W08326 are mutated Tet repressor (TetR) proteins. They were  
 CC used to create fusion protein, comprising a mutated tetR fused to a  
 CC transcriptional control sequence (e.g. v-erbA silencer domain), to  
 CC produce tetracycline (Tc)-controlled transcriptional activators (ETA).  
 CC The main invention of the specification concerns modified TetR proteins  
 CC that bind to modified class B tet operator sequences tetO-4C and  
 CC tetO-6C (see T45711 and T48478). Modified TetR proteins can be fused  
 CC to any transcription regulatory polypeptide and used to control  
 CC transcription of a tetO-4C or tetO-6C linked gene. Nucleic acid  
 CC encoding such a fusion protein may be introduced into a cell and  
 CC transcription of the protein can be controlled by altering the concn.  
 CC of tetracycline (or an analogue) in the cell, as appropriate. This  
 CC ability to modulate gene expression in a predictable way is very useful  
 CC in gene therapy and for recombinant protein prodn. in cultured cells or  
 CC transgenic animals. The Tc-inducible system is also useful for the  
 CC prodn. of transgenic animal models for the study of disease and also  
 CC for the study of gene function e.g. during differentiation. The  
 CC Tc-inducible system allows rapid activation of gene transcription  
 CC without cellular toxicity, high concns. of inducer are not required.  
 SQ Sequence 207 AA;

Query Match 98.9%; Score 1068; DB 1; Length 207;  
 Best Local Similarity 99.0%; Pred. No. 1.3e-111;  
 Matches 205; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSRLDQSKVINSALLELLNEVGIEGLTRKLAQKLGVEOPTLYWHVKNKRALDALAEML 60  
 DB 1 MSRLDQSKVINSALLELLNEVGIEGLTRKLAQKLGVEOPTLYWHVKNKRALDALAEML 60  
 QY 61 DRHHTFCPLGESWQDFLRNKAQKSFRCALLSHRDGAKVHLGTRPTEKQYETLENQALFL 120  
 DB 61 DRHHTFCPLGESWQDFLRNKAQKSFRCALLSHRDGAKVHLGTRPTEKQYETLENQALFL 120  
 QY 121 CQGFSLNALLYALSAGVHFTLCGVLEDOHQVAKEREPTTDSMPPLRQALIEFDHQ 180  
 DB 121 CQGFSLNALLYALSAGVHFTLCGVLEDOHQVAKEREPTTDSMPPLRQALIEFDHQ 180  
 QY 181 GAEPALFGLLEIICGLEKQKCESGS 207  
 DB 181 GAEPALFGLLEIICGLEKQKCESGS 207

Search completed: June 9, 1999, 12:59:17  
 Job time: 9196 sec

Fri Jun 11 11:58:05 1999

us-08-486-814-17.rag

---

Page 8



GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 9, 1999, 10:27:01 ; Search time 19.42 Seconds  
(without alignments)  
399.294 Million cell updates/sec

Title: US-08-486-814-17  
Perfect score: 1080  
Sequence: 1 MSRLDKSVINSALLENEV.....FGLELIICGLEKQLKCESGS 207  
Scoring table: BLOSUM62

Searched: 116738 seqs, 37460341 residues

Database : PIR\_58:.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1075	99.5	207	1 RPCTN	repressor tetr - E
2	683	63.2	218	2 S07359	regulatory protein
3	683	63.2	218	2 S30287	regulatory protein
4	619	57.3	126	2 A26948	tetracycline resis
5	525	48.6	219	1 RPECYS	tetracycline repre
6	515.5	47.7	216	2 JQ1478	regulatory protein
7	515.5	47.7	216	2 S38655	tetr protein - Pse
8	509.5	47.2	211	2 I77569	tet repressor - Es
9	505.5	46.8	216	1 RPECR1	tet repressor prot
10	176.5	16.3	261	2 H70740	hypothetical prote
11	149.5	13.8	259	2 A40046	Tetr repressor hom
12	119	11.0	191	2 A69900	hypothetical prote
13	110.5	10.2	210	2 S42417	probable transcrip
14	110.5	10.2	210	2 S42419	probable transcrip
15	108.5	10.0	189	2 H70042	transcription regu
16	101.5	9.4	192	2 C70487	transcription regu
17	98	9.1	196	2 B70827	hypothetical prote
18	95.5	8.8	196	2 D64918	glucuronide repres
19	94.5	8.8	236	2 S75298	hypothetical prote
20	94	8.7	179	2 B70391	transcription regu
21	93.5	8.7	246	2 E70861	hypothetical prote
22	92	8.5	213	2 F70946	probable regulato
23	89	8.2	194	2 C70035	transcription regu
24	87.5	8.1	235	2 A64695	conserved hypothet
25	86.5	8.0	192	2 E69280	iron-dependent rep
26	86	8.0	205	2 S29308	hypothetical prote
27	85.5	7.9	236	2 D64855	hypothetical prote
28	85.5	7.9	677	2 I49045	SKD3 - mouse
29	85	7.9	198	2 E69779	transcription regu
30	84	7.8	188	2 S12393	hypothetical prote
31	84	7.8	207	2 H70001	hypothetical prote
32	83.5	7.7	216	2 C70649	nitrogen fixation
33	83	7.7	519	2 S31136	regulatory protein
34	83	7.7	195	2 S10899	Tcd37 protein - fr
35	83	7.7	365	2 S17885	conserved hypothet
36	83	7.7	192	2 F70487	hypothetical prote
37	82	7.6	260	2 F64995	hypothetical prote
38	82	7.6	629	2 A71023	probable arginyl-t
39	81.5	7.5	299	2 S39744	transcription regu

RESULT 1  
RPECTN  
repressor tetr - Escherichia coli transposon Tn10  
C:Species: Escherichia coli  
C:Date: 17-Mar-1987 #sequence\_revision 17-Mar-1987 #text\_change 12-Jun-1998  
C:Accession: A03576; S02667  
R:Postle, K.; Nguyen, T.T.; Bertrand, K.P.  
Nucleic Acids Res. 12, 4849-4863, 1984  
A:Title: Nucleotide sequence of the repressor gene of the Tn10 tetracycline resistance  
A:Reference number: A03576; MUID:84247342  
A:Accession: A03576  
A:Molecule type: DNA  
A:Residues: 1-207 <POS>  
A:Cross-references: GB:X0694; NID:943051; PID:943052  
R:Altschmid, L.; Baumeister, R.; Pfeleiderer, K.; Hillen, W.  
EMBO J. 7, 4011-4017, 1988  
A:Title: A threonine to alanine exchange at position 40 of tet repressor alters the r  
A:Reference number: S02667; MUID:89091153  
A:Accession: S02667  
A:Status: not compared with conceptual translation  
A:Molecule type: DNA  
A:Residues: 1-100 <ALT>  
A:Cross-references: GB:J01830; NID:gl54845  
C:Comment: This protein contains sequences similar to the DNA recognition regions of  
A:Genetics:  
A:Gene: tetr  
C:Superfamily: tetracycline repressor  
C:Keywords: antibiotic resistance; DNA binding; transcription regulation  
F:25-46/Region: helix-turn-helix #status predicted

Query Match	99.5%	Score 1075;	DB 1;	Length 207;
Best Local Similarity	99.5%;	Pred. No. 2.le-92;		
Matches 206;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;
QY	1	MSRLDKSVINSALLENEVGIEGLTTRKLAQKLGVEQPTLYWHVKNRALLDALAIEML	60	
Db	1	MSRLDKSVINSALLENEVGIEGLTTRKLAQKLGVEQPTLYWHVKNRALLDALAIEML	60	
QY	61	DRHHTHFCPLEGESWDFLRNKAKSFRCALLSHRDGAKVHLGTRTEKQYETLENQALFL	120	
Db	61	DRHHTHFCPLEGESWDFLRNKAKSFRCALLSHRDGAKVHLGTRTEKQYETLENQALFL	120	
QY	121	COQGSLENALYALSAVGHTLGCVCLEQEHQVAKEREETPTTDSMPPLRLQALFQHQ	180	
Db	121	COQGSLENALYALSAVGHTLGCVCLEQEHQVAKEREETPTTDSMPPLRLQALFQHQ	180	
QY	181	GAEPAPFLFGLLEIICGLEKQLKCESGS	207	
Db	181	GAEPAPFLFGLLEIICGLEKQLKCESGS	207	

RESULT 2  
S07359  
regulatory protein tetr - Escherichia coli plasmid RAI  
N:Alternate names: tet repressor class D  
C:Species: Escherichia coli  
C:Date: 07-Sep-1990 #sequence\_revision 07-Sep-1990 #text\_change 12-Jun-1998  
C:Accession: S07359  
R:Unger, B.; Klock, G.; Hillen, W.  
Nucleic Acids Res. 12, 7693-7703, 1984  
A:Title: Nucleotide sequence of the repressor gene of the RAI tetracycline resistance

A:Reference number: S07359; MUID:85037938

A:Accession: S07359

A:Molecule type: DNA

A:Residues: 1-218 <UNG>

A:Cross-references: EMBL:X01083; NID:g43053; PID:g43054

C:Genetics:

A:Gene: tetr

A:Genome: plasmid

C:Superfamily: tetracycline repressor

C:Keywords: antibiotic resistance; DNA binding; transcription regulation

F:25-46/Region: helix-turn-helix #status predicted

Query Match 63.2%; Score 683; DB 2; Length 218;  
Best Local Similarity 65.2%; Pred. No. 4.1e-56;  
Matches 131; Conservative 26; Mismatches 44; Indels 0; Gaps 0;

QY 1 MSRLDKSVINSALLELLNEVGIEGLTRKLAQKLGVEOPTLYWHVKNKRALDLALAEML 60

Db 1 MARLNRESVIDAALLELLNETGIDGLTRKLAQKLGIEOPTLYWHVKNKRALDLALAEVL 60

QY 61 DRHHTFCPLEGESWQDFLNKAKSFRCALLSHRDGAKVHLGTRPTEKQYETLENQALFL 120

Db 61 ARHHDYSLPAAGESWQDFLNKAKSFRALLRYRDGAKVHLGTRPDEKQYDVTETOLREM 120

QY 121 CQOGFSLNALYALSAYGHFTGLGCVLEDOEHQVAKKERETPTTDSMPPLLRQALIEFDHQ 180

Db 121 TENGFSLRDLGLYSAVSHTLGNVLEQOEHTAALTDRAAPDENLPLLRALQIMQSD 180

QY 181 GAEPAFLEGLICGLEKOL 201

Db 181 DGEQAFLEGLIESLRGFEVOL 201

RESULT 3

S30287

regulatory protein tetr - Salmonella ordonez plasmid pip173

C:Species: Salmonella ordonez

C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 12-Jun-1998

C:Accession: S30287

R:Allard, J.D.; Gibson, M.L.; Vu, L.H.; Nguyen, T.T.; Bertrand, K.P.

Mol. Gen. Genet. 237, 301-305, 1993

A:Title: Nucleotide sequence of class D tetracycline resistance genes from Salmonella or

A:Reference number: S30286

A:Molecule type: DNA

A:Residues: 1-218 <ALL>

A:Cross-references: EMBL:X65876; NID:g49073; PID:g49075

C:Genetics:

A:Gene: tetr

A:Genome: plasmid

C:Superfamily: tetracycline repressor

C:Keywords: antibiotic resistance; DNA binding; transcription regulation

F:25-46/Region: helix-turn-helix #status predicted

Query Match 63.2%; Score 683; DB 2; Length 218;  
Best Local Similarity 65.2%; Pred. No. 4.1e-56;  
Matches 131; Conservative 26; Mismatches 44; Indels 0; Gaps 0;

QY 1 MSRLDKSVINSALLELLNEVGIEGLTRKLAQKLGVEOPTLYWHVKNKRALDLALAEML 60

Db 1 MARLNRESVIDAALLELLNETGIDGLTRKLAQKLGIEOPTLYWHVKNKRALDLALAEVL 60

QY 61 DRHHTFCPLEGESWQDFLNKAKSFRCALLSHRDGAKVHLGTRPTEKQYETLENQALFL 120

Db 61 ARHHDYSLPAAGESWQDFLNKAKSFRALLRYRDGAKVHLGTRPDEKQYDVTETOLREM 120

QY 121 CQOGFSLNALYALSAYGHFTGLGCVLEDOEHQVAKKERETPTTDSMPPLLRQALIEFDHQ 180

Db 121 TENGFSLRDLGLYSAVSHTLGNVLEQOEHTAALTDRAAPDENLPLLRALQIMQSD 180

QY 181 GAEPAFLEGLICGLEKOL 201

Db 181 DGEQAFLEGLIESLRGFEVOL 201

RESULT 4

A26948

tetracycline resistance protein - Haemophilus parainfluenzae

C:Species: Haemophilus parainfluenzae

C:Date: 19-Nov-1988 #sequence\_revision 19-Nov-1988 #text\_change 12-Jun-1998

C:Accession: A26948

R:Heuer, C.; Hickman, R.K.; Curiale, M.S.; Hillen, W.; Levy, S.B.

J. Bacteriol. 169, 990-994, 1987

A:Title: Constitutive expression of tetracycline resistance mediated by a Tn10-like e

A:Reference number: A26948; MUID:87137315

A:Accession: A26948

A:Molecule type: DNA

A:Residues: 1-126 <HEU>

A:Cross-references: GB:M15539; NID:g148988; PID:g148989

C:Genetics:

A:Gene: tetr

C:Superfamily: tetracycline repressor

F:25-46/Region: helix-turn-helix #status predicted

Query Match 57.3%; Score 619; DB 2; Length 126;  
Best Local Similarity 95.2%; Pred. No. 1.7e-50;  
Matches 119; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 MSRLDKSVINSALLELLNEVGIEGLTRKLAQKLGVEOPTLYWHVKNKRALDLALAEML 60

Db 1 MSRLDKSVINSALLELLNEVGIEGLTRKLAQKLGVEOPTLYWHVKNKRALDLALAEML 60

QY 61 DRHHTFCPLEGESWQDFLNKAKSFRCALLSHRDGAKVHLGTRPTEKQYETLENQALFL 120

Db 61 DRHHTFCPLEGESWQDFLNKAKSFRCALLSHRDGAKVHLGTRPTEKQYETLENQALFY 120

QY 121 CQOGF 125

Db 121 ANKVF 125

RESULT 5

RPECYS

tetracycline repressor - Escherichia coli plasmid pSC101

C:Species: Escherichia coli

C:Date: 28-Aug-1985 #sequence\_revision 28-Aug-1985 #text\_change 12-Jun-1998

C:Accession: A03575

R:Brow, M.A.D.; Pesin, R.; Sutcliffe, J.G.

Mol. Biol. Evol. 2, 1-12, 1985

A:Title: The tetracycline repressor of pSC101.

A:Reference number: A03575; MUID:88216101

A:Accession: A03575

A:Molecule type: DNA

A:Residues: 1-219 <BRO>

A:Cross-references: GB:M36272; NID:g150945; PID:g150946

C:Comment: This protein is the repressor of the tetracycline resistance element; its

C:Genetics:

A:Gene: tetr

A:Genome: plasmid

C:Superfamily: tetracycline repressor

C:Keywords: antibiotic resistance; DNA binding; transcription regulation

F:24-46/Domain: DNA binding #status predicted <DBN>

F:25-46/Region: helix-turn-helix #status predicted

Query Match 48.6%; Score 525; DB 1; Length 219;

Best Local Similarity 52.7%; Pred. No. 1.7e-41;

Matches 107; Conservative 32; Mismatches 60; Indels 4; Gaps 1;

QY 1 MSRLDKSVINSALLELLNEVGIEGLTRKLAQKLGVEOPTLYWHVKNKRALDLALAEML 60

Db 1 MNKLOEAVIRTALELLNDVGMGLTRRLAERLGVQOPALYWHFKNKRALDLALAEML 60



tet repressor protein (Tn 1721) - Escherichia coli plasmid RP1  
 C:Species: Escherichia coli  
 C:Date: 25-Feb-1985 #sequence\_revision 25-Feb-1985 #text\_change 12-Jun-1998  
 C:Accession: A03574  
 R:Waters, S.H.; Rogowsky, P.; Grinstead, J.; Altenbuchner, J.; Schmitt, R.  
 Nucleic Acids Res. 11, 6089-6105, 1983  
 A:Title: The tetracycline resistance determinants of RP1 and Tn1721: nucleotide sequence  
 A:Reference number: A93486; MUID:83299270  
 A:Accession: A03574  
 A:Molecule type: DNA  
 A:Residues: 1-216 <WAT>  
 A:Cross-references: GB:X00006; NID:g42508; PID:g42509  
 C:Genetics:  
 A:Gene: tetR  
 A:Genome: plasmid  
 C:Superfamily: tetracycline repressor  
 C:Keywords: antibiotic resistance; DNA binding; transcription regulation  
 F:25-46/Region: helix-turn-helix #status predicted

Query Match 46.8%; Score 505.5; DB 1; Length 216;  
 Best Local Similarity 51.5%; Pred. No. 1.1e-39;  
 Matches 104; Conservative 34; Mismatches 61; Indels 3; Gaps 2;

QY 1 MSRLSKVINSALLENEVGIEGTTTRKLAQKLVGEQPTLYWHVKNKRALDIAIEML 60  
 Db 1 MTKLQPTVIRAAALLNEVGVDGLTTRKLAERLGVQPPALYWHFNRKRALDIAEAML 60  
 QY 61 DRHHTHFCPLEGESWQDFLRNKAISFRCALLSHRDGAKVHLGTRPTKEQYETLENQ 120  
 Db 61 AENHSTVPRADDDWRSFELTGNARFQALLAYRDGARIHAGTRPGAPOMETADAQLRFL 120  
 QY 121 COQGSFLENALYSAGVHFTLGCVLDEQEHQVAKEREETPTDTSMP--PLLROAIELED 178  
 Db 121 CEAGSFAGDAVNALMTISYFTVGAVLEQAGDSSEGER-GGTVEQAPLSPLLRAAIDAFD 179  
 QY 179 HOGAEPALFGLLELICGLEKQ 200  
 Db 180 EAGPDNAEQGLAIVDGLAKR 201

RESULT 10  
 H70740  
 Hypothetical protein Rv1353c - Mycobacterium tuberculosis (strain H37Rv)  
 C:Species: Mycobacterium tuberculosis  
 C:Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 17-Jul-1998  
 C:Accession: H70740  
 R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Nature 393, 537-544, 1998  
 A:Authors: Krogh, A.; McLean, J.; Moule, S.; Murphy, L.; Oliver, S.; Osborne, J.; Quail, J.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
 A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
 A:Reference number: A70500; MUID:98295987  
 A:Accession: H70740  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-261 <COL>  
 A:Cross-references: GB:275555; GB:AL123456; NID:g3261608; PID:e250355; PID:g1419056  
 A:Experimental source: strain H37Rv  
 C:Genetics:  
 A:Gene: Rv1353c

Query Match 16.3%; Score 176.5; DB 2; Length 261;  
 Best Local Similarity 28.1%; Pred. No. 3.7e-09;  
 Matches 61; Conservative 31; Mismatches 88; Indels 37; Gaps 5;

QY 4 LDKSKVINSALLENEVGIEGTTTRKLAQKLVGEQPTLYWHVKNKRALDIAIEMLRH 63  
 Db 16 INPEDIISGAFELAQQVSDINLSMPLGLHGVGVTSITVYFRKDDLLNAMTDRALSKY 75  
 QY 64 HTFPCPLEGESWQDFLRNKAISFR-----CALLSHRDGAKVHLGTRPTKEQYETLE 114

Db 76 VFATPYTEAGDWRETLNRHARSMRKTFADNPVLCDDLILR-----AALSPKTAIRLGAQEME 131  
 QY 115 NQLAFLCQGFSLLENALYSAGVHFTLGCVLDEQ-----EHOVAKEERE 159  
 Db 132 KAIANLVMTAGLSUEDAFDIYSYVSVHVRGSLVLDRLSRKSQSAGSGPSAIEHPVAIDPAT 191  
 QY 160 TPTTDSMPPLLRQAIELFDHOGA--EPAPFLFGLLEII 194  
 Db 192 T-----PLLAHATGRGHGRIAGPDETINFEYGLCIL 221

RESULT 11  
 A40046  
 Tet repressor homolog actII-1 - Streptomyces coelicolor  
 C:Species: Streptomyces coelicolor  
 C:Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 12-Jun-1998  
 C:Accession: A40046  
 R:Fernandez-Moreno, M.A.; Caballero, J.L.; Hopwood, D.A.; Malpartida, F.  
 Cell 66, 769-780, 1991  
 A:Title: The act cluster contains regulatory and antibiotic export genes, direct target of the act promoter  
 A:Reference number: A40046; MUID:91347376  
 A:Accession: A40046  
 A:Molecule type: DNA  
 A:Residues: 1-259 <PER>  
 A:Cross-references: GB:M64683; NID:g153143; PID:g455360  
 C:Genetics:  
 A:Start codon: GTG  
 F:52-73/Region: helix-turn-helix #status predicted

Query Match 13.8%; Score 149.5; DB 2; Length 259;  
 Best Local Similarity 26.0%; Pred. No. 1.1e-06;  
 Matches 57; Conservative 35; Mismatches 90; Indels 37; Gaps 8;

QY 4 LDKSKVINSALLENEVGIEGTTTRKLAQKLVGEQPTLYWHVKNKRALDIAIEMLRH 116  
 Db 31 LQDRIIVTALGIDAEGLDLSMRRLAQLKTHGASHYAHVGNRDELDFVDFIVLTV 90  
 QY 58 EMLDRHHTHFCPLEGESWQDFLRNKAISFRCALLSHRDGAKVHLGTRPT-KEQYETLENQ 116  
 Db 91 EVPE-----PEPGR-WAEQVKEMCRSLRMFLAHRDLARIAIDRVPLGPNMGVGHRT 142  
 QY 117 LAFLCQGFSLLENALYSAGVHFTLGCVLDEQEHQVAKERE-----RET 160  
 Db 143 MNLRSGLHDELAAAYGGDLLSTFTVALEQSSRNPTGEQAGVFADQLHGYLKSL 202  
 QY 161 PTTDSMPPLLRQA--IELFDHOGAEPAPFLFGLLEII 197  
 Db 203 PAT-SFPNVLHAGPIISLD---SDRRFELGLEIIAGL 237

RESULT 12  
 A69900  
 Hypothetical protein yobS - Bacillus subtilis  
 C:Species: Bacillus subtilis  
 C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 24-Sep-1998  
 C:Accession: A69900  
 R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Ber  
 C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.;  
 A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari,  
 Nature 390, 249-256, 1997  
 A:Authors: Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galleron, N.; Gh  
 wood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.; Itaya, M  
 Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardiniois, S.; Laube  
 A:Authors: Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mael, C.; Med  
 , K.; Ogilwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle, D.; Por  
 Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon, E.; Sch  
 A:Authors: Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seror, S.J.; Ser  
 amakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama, S.; Vanden  
 ; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.; Yoshikaw  
 A:Authors: Zumstein, E.; Yoshikawa, H.; Danchin, A.  
 A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtili

Db 155 DDLKRETAIVF 165

RESULT 14

S42419

probable transcription repressor mtrR - Neisseria gonorrhoeae (strain FA19)

C:Species: Neisseria gonorrhoeae

A:Variety: strain FA19

C:Date: 07-May-1998 #sequence\_revision 15-May-1998 #text\_change 18-Sep-1998

A:Accession: S42419; S40250

R:Pan, W.; Spratt, B.G.

Mol. Microbiol. 11, 769-775, 1994

A:Title: Regulation of the permeability of the gonococcal cell envelope by the mtr sy

A:Reference number: S42417

A:Accession: S42419

A:Molecule type: DNA

A:Status: nucleic acid sequence not shown

A:Residues: 1-210 <PAN>

A:Cross-references: EMBL:Z25797; NID:g452332; PID:g438189

A:Experimental source: strain FA19

C:Genetics:

A:Gene: mtrR

C:Superfamily: probable transcription repressor mtrR

C:Keywords: DNA binding; transcription regulation

Query Match 10.2%; Score 110.5; DB 2; Length 210;

Best Local Similarity 27.4%; Pred. No. 0.0035;

Matches 52; Conservative 22; Mismatches 73; Indels 43; Gaps 8;

Qy 6 KSKVINSALLENEVGIEGLTRKLAQKLGVEQPTLYWHVKNKRALDLALAIEMLDHRHT 65

Db 12 KEHLMLAALETFYKKGARTSLNEIAQAAGVTRGALYWHFKNEDLFDALFORICDDIEN 71

Qy 66 ----HFCPLEGESWQDFLRNKAISFRCALLSHRDGAKVHLGTRPTEKQYETLENQALFLC 121

Db 72 CIAQDAADAEGGSWTF-RHTLLHFFERLOSNDIHKFH-----NIFLKC 116

Qy 122 CQGSFLENALYALSVAHGFTLGCVLDEQHOVAKEREPTTDSMPPLRQAIE---LFD 178

Db 117 EH--TEQNA--AVIAIA-----RKHQAIWREKIT-----AVLTEAVENQDLAD 155

Qy 179 HOGAEPAPFLF 188

Db 156 DLDKETAVIF 165

RESULT 15

H70042

transcription regulator TetR/AcrR family homolog yvxB - Bacillus subtilis

C:Species: Bacillus subtilis

C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 24-Sep-1998

A:Accession: H70042

R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Ber

C.: Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.;

A.: Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari,

Nature 390, 249-256, 1997

A:Authors: Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galleron, N.; Gh

wood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.; Iteya, M

Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois, S.; Laube

A:Authors: Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel, C.; Mer

, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle, D.; Por

Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanlon, E.; Schl

amakers, Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Sero, S.J.; Ser

A:Authors: Schroeter, R.; Scoffone, F.; Tognoni, A.; Tosato, V.; Uchiyama, S.; Vanden

amashiro, A.; Tanaka, T.; Terpstra, P.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.; Yoshikaw

; Wipac, A.; Yamamoto, H.; Yamane, K.; Danchin, A.

A:Authors: Zumbstein, E.; Yoshikawa, H.; Danchin, A.

A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtili

A:Reference number: A69580; MUID:98044033

A:Accession: H70042

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-189 <KUN>

Db 155 DDLKRETAIVF 165

Query Match 11.0%; Score 119; DB 2; Length 191;

Best Local Similarity 23.6%; Pred. No. 0.0051;

Matches 48; Conservative 39; Mismatches 74; Indels 42; Gaps 8;

Qy 4 LDKSKVINSALLENEVGIEGLTRKLAQKLGVEQPTLYWHVKNKRALDLALAI-----EM 59

Db 7 LTKQWIDAAAEIADQGVNGVSLAASKKMNVRPPSLYHNGHQAIKRAELAVRGLTKL 66

Qy 60 LDRHHTFCPLEGES-----WQDFLRNKAISFRCALLSHRDGAKVHLGTRPTEKQYE 111

Db 67 FQWADSVTERKGDUSAMLSAHAYVDFAIENPGYENALL-----KVH-----DKRTE 114

Qy 112 TLENQALFLCQGSFLENALYALSVAHGFTLG--CVLEQHOVAKEREPTTDSMP 167

Db 115 IVSDQIVCLVTK-LLENGYASEKTAHATRGLRSLRHGFTVLIKAERAFEREEDILESLS 173

Qy 168 PLLRQAIELFDHQAEPAPFLFL 190

Db 174 FSIR-----TFLSGL 183

RESULT 13

S42417

probable transcription repressor mtrR - Neisseria gonorrhoeae (strain CH95)

C:Species: Neisseria gonorrhoeae

A:Variety: strain CH95

C:Date: 07-May-1998 #sequence\_revision 15-May-1998 #text\_change 18-Sep-1998

A:Accession: S42417; S40251

R:Pan, W.; Spratt, B.G.

Mol. Microbiol. 11, 769-775, 1994

A:Title: Regulation of the permeability of the gonococcal cell envelope by the mtr sy

A:Reference number: S42417

A:Accession: S42417

A:Molecule type: DNA

A:Status: nucleic acid sequence not shown

A:Residues: 1-210 <PAN>

A:Cross-references: EMBL:Z25796; NID:g438190; PID:g438191

A:Experimental source: CH95

C:Genetics:

A:Gene: mtrR

C:Superfamily: probable transcription repressor mtrR

C:Keywords: DNA binding; transcription regulation

Query Match 10.2%; Score 110.5; DB 2; Length 210;

Best Local Similarity 27.2%; Pred. No. 0.0035;

Matches 52; Conservative 21; Mismatches 73; Indels 45; Gaps 9;

Qy 6 KSKVINSALLENEVGIEGLTRKLAQKLGVEQPTLYWHVKNKRALDLALAIEMLDHRHT 65

Db 12 KEHLMLAALETFYKKGARTSLNEIAQAAGVTRGALYWHFKNEDLFDALFORICDDIEN 71

Qy 66 ----HFCPLEGESWQDFLRNKAISFRCALLSHRDGAKVHLGTRPTEKQYETLENQALFLC 120

Db 72 CIAQDAADAEGGSW-----TVFRHTLL-----HFFERLOSNDIHKFHILFLK 115

Qy 121 CQGSFLENALYALSVAHGFTLGCVLDEQHOVAKEREPTTDSMPPLRQAIE---LF 177

Db 116 CEH--TEQNA--AVIAIA-----RKHQAIWREKIT-----AVLTEAVENQDLA 154

Qy 178 DHQAEPAPFLF 188

A:Cross-references: GB:Z99122; GB:AL009126; NID:g2636029; PID:e1184426; PID:g2636046  
A:Experimental source: strain 168  
C:Genetics:  
A:Gene: yvkB

```

Query Match      10.0%  Score 108.5;  DB 2;  Length 189;
Best Local Similarity 19.6%  Pred. No. 0.0047;
Matches 41;  Conservative 32;  Mismatches 51;  Indels 85;  Gaps 5;

QY  3 RLDKSKVINSALELLENEVGIEGLTRKLAQKLGVEQPTLYWHVKNKRALIDALAIEMLDL 62
   | : : : : | : | : : : | : | : : : | : | : : : | : | : : : |
Db  2 RPTNKRILDAAMQLLVKKGYRATTKEIAEKANVSEATIFRNFKNKQGLVEALLSHSSN 51
   | : : : : | : | : : : | : | : : : | : | : : : | : | : : : |
QY  63 HHTHFCPLLEGESWODEFLRNKAKSFRCALLSHRDGAKVHLGTRPTEKQYETLENQALFLCQ 122
   | : : : : | : | : : : | : | : : : | : | : : : | : | : : : |
Db  62 RGSILEQTEGDLKDLL-----HIGT----- 83

QY  123 QGFSLENALYALSAVGHFTLGCVLEDOEHO-----VAKEERETPTTDSM 166
   | : : : : | : | : : : | : | : : : | : | : : : | : | : : : |
Db  83 -----CLLELEHRKDIKISFREPA MFQDVINHVIETYP--OSM 119

QY  167 PPLLQRAIELFDHOG-----AEPALFL 188
   | : : : : | : | : : : | : | : : : | : | : : : | : | : : : |
Db  120 KQLLYDYLTGMEKGVITGCTGNEAEHADVF 148

```

Search completed: June 9, 1999, 12:59:50  
Job time: 9169 sec

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 9, 1999, 10:29:01 ; Search time 16.52 Seconds  
(without alignments)  
336.316 Million cell updates/sec

Title: US-08-486-814-17  
Perfect score: 1080  
Sequence: 1 MSRLDKSKVINSALELLNEV.....FGLELICGLEKOLKCESGS 207

Scoring table: BLOSUM62

Searched: 74019 seqs, 26840295 residues

Database : SwissProt\_36:\*

pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1075	99.5	207	1	TER2_ECOLI	P04483 escherichia
2	683	63.2	218	1	TER8_PASPI	P51562 pasteurella
3	678	62.8	217	1	TER4_ECOLI	P09164 escherichia
4	673	62.3	207	1	TER8_PASMU	P51561 pasteurella
5	525	48.6	219	1	TER3_ECOLI	P03039 escherichia
6	515.5	47.7	216	1	TER1_ECOLI	P03038 escherichia
7	509.5	47.2	211	1	TER5_ECOLI	P21337 escherichia
8	487	45.1	210	1	TER7_VIBAN	P51560 vibrio angu
9	176.5	16.3	261	1	Y07H_MYCTO	Q11023 mycobacteri
10	110.5	10.2	210	1	MTRR_NEIGO	P39897 neisseria g
11	95.5	8.8	196	1	UDR_ECOLI	Q59431 escherichia
12	85.5	7.9	677	1	SKD3_MOUSE	Q60649 mus musculu
13	85.5	7.9	236	1	YCFQ_ECOLI	P75952 escherichia
14	84	7.8	188	1	YPC2_STAAU	P23217 staphylococ
15	83	7.7	195	1	BET1_ECOLI	P17446 escherichia
16	83	7.7	519	1	NIFL_AZOVI	P30663 azotobacter
17	83	7.7	955	1	PHSL_IPOBA	P27598 ipomoea bat
18	82	7.6	260	1	YFAX_ECOLI	P77732 escherichia
19	81.5	7.5	299	1	YWFK_BACSU	P39647 bacillus su
20	80.5	7.5	577	1	THR1_SCHPO	Q09684 schizosacch
21	80	7.4	2133	1	FAB_PIG	P12263 sus scrofa
22	80	7.4	212	1	TTK_ECOLI	P06969 escherichia
23	79	7.3	192	1	BM3R_BACME	P43506 bacillus me
24	79	7.3	866	1	MISP_SCHJA	Q05870 schistosoma
25	78.5	7.3	782	1	VACB_HAEIN	P44907 haemophilus
26	78	7.2	215	1	ACRR_ECOLI	P34000 escherichia
27	77.5	7.2	2472	1	SPCN_HUMAN	Q13813 homo sapien
28	77.5	7.2	1448	1	TRG5_ECOLI	P27190 escherichia
29	77.5	7.2	212	1	YCDC_ECOLI	P75899 escherichia
30	77	7.1	2319	1	FAB_MOUSE	Q06194 mus musculu
31	77	7.1	1957	1	YD85_SCHPO	Q10411 schizosacch
32	76.5	7.1	220	1	ENVR_ECOLI	P31676 escherichia
33	76.5	7.1	728	1	HS9A_CHICK	P11501 gallus gall
34	76.5	7.1	963	1	KINE_HUMAN	P33176 homo sapien
35	76.5	7.1	1208	1	YDH6_SCHPO	Q92351 schizosacch
36	76	7.0	218	1	YTK_HAEIN	P29280 haemophilus
37	75	7.0	187	1	Y893_HAEIN	P44923 haemophilus
38	75.5	7.0	549	1	CF1A_DROME	P16241 drosophila
39	75.5	7.0	723	1	HS9B_HUMAN	P08238 homo sapien
40	75.5	7.0	866	1	MISP_SCHMA	P06198 schistosoma
41	75.5	7.0	1391	1	N157_YEAST	P40064 saccharomyc
42	75.5	7.0	888	1	UFO_MOUSE	Q00993 mus musculu
43	74.5	6.9	557	1	ATKA_ECOLI	P03959 escherichia

## ALIGNMENTS

```

RESULT 1
TER2_ECOLI
ID TER2_ECOLI STANDARD: PRT; 207 AA.
AC P04483;
DT 13-AUG-1987 (REL. 05, CREATED)
DT 13-AUG-1987 (REL. 05, LAST SEQUENCE UPDATE)
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE TETRACYCLINE REPRESSOR PROTEIN CLASS B (TRANSPONSON TN10).
GN TETR.
OS ESCHERICHIA COLI.
OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;
OC ENTEROBACTERIACEAE.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 84247342.
RA POSTLE K., NGUYEN T.T., BERTRAND K.P.;
RL NUCLEIC ACIDS RES. 12:4849-4863(1984).
RN [2]
RP SEQUENCE OF 1-22 FROM N.A.
RX MEDLINE; 84005886.
RA BERTRAND K.P., POSTLE K., WRAY L.V. JR., REZNIKOFF W.S.;
RL GENE 23:149-156(1983).
RN [3]
RP MUTAGENESIS.
RX MEDLINE; 89094871.
RA SMITH L.D., BERTRAND K.P.;
J. MOL. BIOL. 203:949-955(1988).
CC -1- FUNCTION: TETR IS THE REPRESSOR OF THE TETRACYCLINE RESISTANCE
ELEMENT; ITS AMINO-TERMINAL REGION FORMS A HELIX-TURN-HELIX
STRUCTURE AND BINDS DNA. BINDING OF TETRACYCLINE TO TETR REDUCES
THE REPRESSOR AFFINITY FOR THE TETRACYCLINE RESISTANCE GENE
(TETA) PROMOTER OPERATOR SITES.
CC -1- SIMILARITY: BELONGS TO THE TETR/ACR FAMILY OF TRANSCRIPTIONAL
REGULATORS.
CC EMBL; X00694; G43052; -.
CC EMBL; J01830; G154846; -.
DR PIR; A03576; RPEPTN.
DR PROSITE; PS01081; HTH_TETR_FAMILY; 1.
KW DNA-BINDING; ANTIBIOTIC RESISTANCE;
FT DNA_BIND 26 45 H-T-H MOTIF (POTENTIAL).
FT SITE 64 64 INVOLVED IN BINDING TO [MG-TC]+ (BY
SIMILARITY).
FT METAL 100 100 MAGNESIUM (OF [MG-TC]+ COMPLEX) (BY
SIMILARITY).
FT MUTAGEN 64 64 H->Y: REDUCES AFFINITY FOR TETRACYCLINE
1000 FOLD..
FT MUTAGEN 82 82 N->H: REDUCES AFFINITY FOR TETRACYCLINE
100 FOLD.
FT MUTAGEN 103 103 T->I: REDUCES AFFINITY FOR TETRACYCLINE
LESS THAN 10 FOLD.
SQ SEQUENCE 207 AA; 23354 MW; 4AB4DD5A CRC32;

```

Query Match 99.5%; Score 1075; DB 1; Length 207;  
Best Local Similarity 99.5%; pred. No. 5.6e-93;  
Matches 206; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

Oy 1 MSRLDKSKVINSALELLNEVGTEGLTTRKLAQKLGVEQFTLYWHVKNRALLDALAIEML 60
    |||||
Db 1 MSRLDKSKVINSALELLNEVGTEGLTTRKLAQKLGVEQFTLYWHVKNRALLDALAIEML 60
    |||||
Oy 61 DRHHTHFCPLEGESWQDFLRNNAKSKFRCALLSHRDGAKVHLGTRPTEKQYETLENQIAFL 120
    |||||
Db 61 DRHHTHFCPLEGESWQDFLRNNAKSKFRCALLSHRDGAKVHLGTRPTEKQYETLENQIAFL 120
    |||||

```

QY 121 CQGFSLNLYALSVAHGHTLGCVLDEQHOVAKEREETPTTDSMPPLLRQALFHDQ 180  
 DB 121 CQGFSLNLYALSVAHGHTLGCVLDEQHOVAKEREETPTTDSMPPLLRQALFHDQ 180  
 QY 181 GAEPFLGLELIICGLEKQKCSGS 207  
 DB 181 GAEPFLGLELIICGLEKQKCSGS 207

RESULT 2  
 TER8\_PASPI  
 ID TER8\_PASPI STANDARD; PRT; 218 AA.  
 AC PS1562;  
 DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)  
 DI 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)  
 DE 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)  
 DN TETRACYCLINE REPRESSOR PROTEIN CLASS H.  
 GN TETR.  
 OS PASTURELLA PISCICIDA.  
 OG PLASMID PSP9351.  
 OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;  
 CC PASTURELLACEAE.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 94328963.  
 RA KIM E.H., AOKI T.;  
 RL MICROBIOL. IMMUNOL. 38:31-38(1994).  
 CC -!- FUNCTION: TETR IS THE REPRESSOR OF THE TETRACYCLINE RESISTANCE  
 CC ELEMENT; ITS AMINO-TERMINAL REGION FORMS A HELIX-TURN-HELIX  
 CC STRUCTURE AND BINDS DNA. BINDING OF TETRACYCLINE TO TETR REDUCES  
 CC THE REPRESSOR AFFINITY FOR THE TETRACYCLINE RESISTANCE GENE  
 CC (TETA) PROMOTER OPERATOR SITES.  
 CC -!- INDUCTION: BY THE [MG-TETRACYCLINE]+ COMPLEX.  
 CC -!- SIMILARITY: BELONGS TO THE TETR/ACRR FAMILY OF TRANSCRIPTIONAL  
 CC REGULATORS.  
 CC EMBL; D16172; G575938; -.  
 DR PROSITE; PS01081; HTH\_TETR\_FAMILY; 1.  
 KW TRANSCRIPTION REGULATION; REPRESSOR; DNA-BINDING;  
 KW ANTIBIOTIC RESISTANCE; PLASMID.  
 FT DNA\_BIND 26 45 H-T-H MOTIF (POTENTIAL).  
 FT SITE 64 64 INVOLVED IN BINDING TO [MG-TC]+ (BY  
 FT METAL 100 100 MAGNESIUM (OF [MG-TC]+ COMPLEX) (BY  
 FT FT SIMILARITY).  
 FT SEQUENCE 218 AA; 24419 MW; 3C57D4AB CRC32;

Query Match 63.2%; Score 683; DB 1; Length 218;  
 Best Local Similarity 65.2%; Pred. No. 1.5e-56;  
 Matches 131; Conservative 26; Mismatches 44; Indels 0; Gaps 0;

QY 1 MSRLDQSKVINSALLELLNEVGIEGLTTRKLAQKLGVEQPTLYWHVKNKRALDALAEML 60  
 DB 1 MARLNRESVIDAALLNNEVGIEGLTTRKLAQKLGVEQPTLYWHVKNKRALDALAEML 60  
 QY 61 DRHHTHFCPLEGESWQDFLRNKAQKFRALLSHRDGAKVHLGTRPTKEQYETLENQALFL 120  
 DB 61 ARHHDYSLPAAGESWQDFLRNKAQKFRALLSHRDGAKVHLGTRPTKEQYETLENQALFL 120  
 QY 121 CQGFSLNLYALSVAHGHTLGCVLDEQHOVAKEREETPTTDSMPPLLRQALFHDQ 180  
 DB 121 TENGFSRLDGLYSAVSHGHTLGVLEQHOVAKEREETPTTDSMPPLLRQALFHDQ 180  
 QY 181 GAEPFLGLELIICGLEKQK 201  
 DB 181 DGEQFLHGLSLEIRGFEVOL 201

RESULT 3  
 TER4\_ECOLI  
 ID TER4\_ECOLI STANDARD; PRT; 217 AA.  
 AC PS9164;  
 DT 01-MAR-1989 (REL. 10, CREATED)

DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
 DE TETRACYCLINE REPRESSOR PROTEIN CLASS D.  
 GN TETR.  
 OS ESCHERICHIA COLI, AND SALMONELLA ORDONEZ.  
 OG PLASMID PIP173.  
 OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;  
 CC ENTEROBACTERIACEAE.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX SPECIES-E. COLI; STRAIN-J53-1; PLASMID-RA1;  
 RX MEDLINE; 85037938.  
 RA UNGER B., KLOCK G., HILLEN W.;  
 RL NUCLEIC ACIDS RES. 12:7693-7703(1984).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX SPECIES-S. ORDONEZ; STRAIN-BM2000; PLASMID-PIP173;  
 RX MEDLINE; 93204906.  
 RA ALLARD J.D., GIBSON M.L., VU L.H., NGUYEN T.T., BERTRAND K.P.;  
 RL MOL. GEN. GENET. 237:301-305(1993).  
 RN [3]  
 RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS), AND PARTIAL SEQUENCE.  
 RX MEDLINE; 94204640.  
 RA HINRICHS W., KISKER C., DUEVEL C., MUELLER A., TOVAR K., HILLEN W.,  
 RA SAENGER W.;  
 RL SCIENCE 264:418-420(1994).  
 CC -!- FUNCTION: TETR IS THE REPRESSOR OF THE TETRACYCLINE RESISTANCE  
 CC ELEMENT; ITS AMINO-TERMINAL REGION FORMS A HELIX-TURN-HELIX  
 CC STRUCTURE AND BINDS DNA. BINDING OF TETRACYCLINE TO TETR REDUCES  
 CC THE REPRESSOR AFFINITY FOR THE TETRACYCLINE RESISTANCE GENE  
 CC (TETA) PROMOTER OPERATOR SITES.  
 CC -!- SUBUNIT: HOMODIMER.  
 CC -!- INDUCTION: BY THE [MG-TETRACYCLINE]+ COMPLEX.  
 CC -!- SIMILARITY: BELONGS TO THE TETR/ACRR FAMILY OF TRANSCRIPTIONAL  
 CC REGULATORS.  
 CC EMBL; X01083; G43054; -.  
 DR EMBL; X65876; G49075; -.  
 DR PIR; S07359; S07359.  
 DR PIR; S30287; S30287.  
 DR PDB; 2TCT; 03-APR-96.  
 DR PDB; 2TRT; 20-JUN-96.  
 DR PROSITE; PS01081; HTH\_TETR\_FAMILY; 1.  
 KW TRANSCRIPTION REGULATION; REPRESSOR; DNA-BINDING; PLASMID;  
 KW ANTIBIOTIC RESISTANCE; 3D-STRUCTURE.  
 FT INIT\_MET 0 0  
 FT DNA\_BIND 25 44 H-T-H MOTIF (POTENTIAL).  
 FT SITE 63 63 INVOLVED IN BINDING TO [MG-TC]+.  
 FT METAL 99 99 MAGNESIUM (OF [MG-TC]+ COMPLEX).  
 SQ SEQUENCE 217 AA; 24287 MW; 830EF8EB CRC32;

Query Match 62.8%; Score 678; DB 1; Length 217;  
 Best Local Similarity 65.0%; Pred. No. 4.4e-56;  
 Matches 130; Conservative 26; Mismatches 44; Indels 0; Gaps 0;

QY 2 SRLDQSKVINSALLELLNEVGIEGLTTRKLAQKLGVEQPTLYWHVKNKRALDALAEML 61  
 DB 1 ARLNRESVIDAALLNNEVGIEGLTTRKLAQKLGVEQPTLYWHVKNKRALDALAEML 60  
 QY 62 RHHTHFCPLEGESWQDFLRNKAQKFRALLSHRDGAKVHLGTRPTKEQYETLENQALFL 121  
 DB 61 RHHDYSLPAAGESWQDFLRNKAQKFRALLSHRDGAKVHLGTRPTKEQYETLENQALFL 120  
 QY 122 CQGFSLNLYALSVAHGHTLGCVLDEQHOVAKEREETPTTDSMPPLLRQALFHDQ 181  
 DB 122 ENGFSRLDGLYSAVSHGHTLGVLEQHOVAKEREETPTTDSMPPLLRQALFHDQ 181  
 QY 182 AEPFLGLELIICGLEKQK 201  
 DB 181 GEPFLHGLSLEIRGFEVOL 200

RESULT 4



```

TER8_PASMU
ID TER8_PASMU STANDARD; PRT: 207 AA.
AC P51561;
DT 01-OCT-1996 (REL. 34, CREATED)
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DE 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE TETRACYCLINE REPRESSOR PROTEIN CLASS H.
GN TETR.
OS PASTURELLA MULTOCIDA.
OG PLASMID PV111.
OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;
OC PASTURELLACEAE.
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN-P2862;
RX MEDLINE; 94153051.
RA HANSEN L.M., MCMURRY L.M., LEVY S.B., HIRSH D.C.;
RL ANTIMICROB. AGENTS CHEMOTHER. 37:2699-2705(1993).
CC -1- FUNCTION: TETR IS THE REPRESSOR OF THE TETRACYCLINE RESISTANCE
CC ELEMENT; ITS AMINO-TERMINAL REGION FORMS A HELIX-TURN-HELIX
CC STRUCTURE AND BINDS DNA. BINDING OF TETRACYCLINE TO TETR REDUCES
CC THE REPRESSOR AFFINITY FOR THE TETRACYCLINE RESISTANCE GENE
CC (TETR) PROMOTER OPERATOR SITES.
CC -1- INDUCTION: BY THE [MG-TETRACYCLINE]+ COMPLEX.
CC -1- SIMILARITY: BELONGS TO THE TETR/ACRR FAMILY OF TRANSCRIPTIONAL
CC REGULATORS.
DR EMBL; U00792; G392872; -.
DR PROSITE; PS01081; HTH_TETR_FAMILY; 1.
KW TRANSCRIPTION REGULATION; REPRESSOR; DNA-BINDING;
KW ANTI-BIOTIC RESISTANCE; PLASMID.
FT DNA_BIND 26 45 H-T-H MOTIF (POTENTIAL).
FT SITE 64 64 INVOLVED IN BINDING TO [MG-TC]+ (BY
FT METAL 100 100 MAGNESIUM (OF [MG-TC]+ COMPLEX) (BY
FT SEQUENCE 207 AA; 23156 MW; A51DA28B CRC32;

Query Match 62.3%; Score 673; DB 1; Length 207;
Best Local Similarity 64.2%; Pred. No. 1.2e-55;
Matches 131; Conservative 27; Mismatches 44; Indels 2; Gaps 1;

QY 1 MSRLDKSVINSALELLNEVGIEGLTRKLAQKLGVEOPTLYHVHVKRRALLDALAIEML 60
DB 1 MAKLDKEQVIDDALILLNEVGIEGLTRRNAQKIGVEOPTLYHVHVKRRALLDALAETIL 60
QY 61 DRHHTHFCPLEGESWQDFLRNKAQKLSHRDGAQKVLHGTRETEKQYETLENQALFL 120
DB 61 QKHHHVLPNETWQDFLRNKAQKLSHRDGAQKVLHGTRETEKQYETLENQALFL 120
QY 121 COQGSFLENALYASAVGHFTLGCVLQDQEHQVAKEREPTTDSM--PPLLROAIELFD 178
DB 121 CDAGFSLSQAVVALLSIAHFTLGSVLETHQEQSKEREKVEITDVTAYPPLLTQAVAIMD 180
QY 179 HOGAEPALFGLLELICGLEKQLK 202
DB 181 SDNGDAALFVLDVNMISGLETVLK 204

RESULT 5
TER3_ECOLI
ID TER3_ECOLI STANDARD; PRT: 219 AA.
AC P03039;
DT 21-JUL-1986 (REL. 01, CREATED)
DT 21-JUL-1986 (REL. 01, LAST SEQUENCE UPDATE)
DE 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE TETRACYCLINE REPRESSOR PROTEIN CLASS C.
GN TETR.
OS ESCHERICHIA COLI.
OG PLASMID PSC101.
OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;
OC ENTEROBACTERIACEAE.
RN [1]

```

```

RP SEQUENCE FROM N.A.
RX MEDLINE; 88216101.
RA BROW M.A.D., PESIN R., SUTCLIFFE J.G.;
RL MOL. BIOL. EVOL. 2:1-12(1985).
CC -1- FUNCTION: TETR IS THE REPRESSOR OF THE TETRACYCLINE RESISTANCE
CC ELEMENT; ITS AMINO-TERMINAL REGION FORMS A HELIX-TURN-HELIX
CC STRUCTURE AND BINDS DNA. BINDING OF TETRACYCLINE TO TETR REDUCES
CC THE REPRESSOR AFFINITY FOR THE TETRACYCLINE RESISTANCE GENE
CC (TETR) PROMOTER OPERATOR SITES.
CC -1- SIMILARITY: BELONGS TO THE TETR/ACRR FAMILY OF TRANSCRIPTIONAL
CC REGULATORS.
DR EMBL; M36272; G150946; -.
DR PIR; A03575; RPECYS.
DR PROSITE; PS01081; HTH_TETR_FAMILY; 1.
KW TRANSCRIPTION REGULATION; REPRESSOR; DNA-BINDING; PLASMID;
KW ANTI-BIOTIC RESISTANCE.
FT DNA_BIND 26 45 H-T-H MOTIF (POTENTIAL).
FT SITE 64 64 INVOLVED IN BINDING TO [MG-TC]+ (BY
FT METAL 100 100 MAGNESIUM (OF [MG-TC]+ COMPLEX) (BY
FT SEQUENCE 219 AA; 24174 MW; 6C9E1828 CRC32;

Query Match 48.6%; Score 525; DB 1; Length 219;
Best Local Similarity 52.7%; Pred. No. 7.2e-42;
Matches 107; Conservative 32; Mismatches 60; Indels 4; Gaps 1;

QY 1 MSRLDKSVINSALELLNEVGIEGLTRKLAQKLGVEOPTLYHVHVKRRALLDALAIEML 60
DB 1 MNKLQREAVIRTALELLNDVGMGEGTLTRRLAERLGVQOPALYHFKNRALLDALAEAML 60
QY 61 DRHHTHFCPLEGESWQDFLRNKAQKLSHRDGAQKVLHGTRETEKQYETLENQALFL 120
DB 61 TINHTHSTPRDDDDRSFLKGNACFSRRALLAYRDGARIHAGTRPAAPQMEKADQLRFL 120
QY 121 COQGSFLENALYASAVGHFTLGCVLQDQEHQVAKERE---TPTDSMPPLLRQAIEL 176
DB 121 CDAGFSAGDATYALMAISYFTVGAVLEQQAQSEADAEERGEDQLTTSASTMPARLQSAWKI 180
QY 177 FDHQAEPALFGLLELICGLEK 199
DB 181 VVEGGDPAFAFERGLALIIGGLEK 203

RESULT 6
TER1_ECOLI
ID TER1_ECOLI STANDARD; PRT: 216 AA.
AC P03038;
DT 21-JUL-1986 (REL. 01, CREATED)
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE TETRACYCLINE REPRESSOR PROTEIN CLASS A (TRANSPONSON 1721).
GN TETR.
OS ESCHERICHIA COLI.
OG PLASMID RP1, AND PLASMID RP4.
OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;
OC ENTEROBACTERIACEAE.
RN [1]
RN SEQUENCE FROM N.A.
RC TRANSPONSON-TN1721;
RX MEDLINE; 92192465.
RA ALLMEIER H., CRESNAR B., GRECK M., SCHMITT R.;
RL GENE 111:11-20(1992).
RN [2]
RN SEQUENCE FROM N.A.
RC PLASMID-PLASMID RP4;
RC TRUEMAN P., SHARPE G.S., BARTH P.T.;
RA SUBMITTED (NOV-1993) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [3]
RN SEQUENCE FROM N.A.
RC TRANSPONSON-TN1721; PLASMID-RP1;
RX MEDLINE; 83299270.

```

RA WATERS S.H., ROGOWSKY P., GRINSTED J., ALTENBUCHNER J., SCHMITT R.;  
 RL NUCLEIC ACIDS RES. 11:6089-6105(1983).  
 CC -!- FUNCTION: TETR IS THE REPRESSOR OF THE TETRACYCLINE RESISTANCE  
 CC ELEMENT; ITS AMINO-TERMINAL REGION FORMS A HELIX-TURN-HELIX  
 CC STRUCTURE AND BINDS DNA. BINDING OF TETRACYCLINE TO TETR REDUCES  
 CC THE REPRESSOR AFFINITY FOR THE TETRACYCLINE RESISTANCE GENE  
 CC (TETA) PROMOTER OPERATOR SITES.  
 CC -!- INDUCTION: BY THE [MG-TETRACYCLINE]+ COMPLEX.  
 CC -!- SIMILARITY: BELONGS TO THE TETR/ACRR FAMILY OF TRANSCRIPTIONAL  
 CC REGULATORS.  
 CC EMBL: X61367; G48198; -.  
 DR EMBL: X75761; G415985; -.  
 DR EMBL: X00006; G42509; -.  
 DR PIR: A03574; RPECRI.  
 DR PROSITE: PS01081; HTH\_TETR\_FAMILY; 1.  
 KW TRANSCRIPTION REGULATION; REPRESSOR; TRANSPOSABLE ELEMENT;  
 KW DNA-BINDING; ANTIBIOTIC RESISTANCE; PLASMID.  
 FT DNA\_BIND 26 45 H-T-H MOTIF.  
 FT SITE 64 64 INVOLVED IN BINDING TO [MG-TC]+ (BY  
 FT SIMILARITY).  
 FT METAL 100 100 MAGNESIUM (OF [MG-TC]+ COMPLEX) (BY  
 FT CONFLICT 65 66 TH -> ST (IN REF. 3).  
 FT CONFLICT 80 80 I -> T (IN REF. 3).  
 FT CONFLICT 154 155 DA -> ES (IN REF. 3).  
 FT SEQUENCE 216 AA; 23320 MW; F72BDFB1 CRC32;

Query Match 47.7%; Score 515.5; DB 1; Length 216;  
 Best Local Similarity 52.5%; Pred. No. 5.4e-41;  
 Matches 106; Conservative 33; Mismatches 60; Indels 3; Gaps 2;

QY 1 MSRLDKSVINSALELLNEVGIEGLTRKLAOKLGVQPTLYWHVKNRALLDALAIEM 60  
 DB 1 MTKLPNTVIRALDNLNEVGVDGLTRKLAERLGVQOPALYHFRNKRALLDALAEAM 60  
 QY 61 DRHHTHFCPLGESQWDFLRNKAISFRCALLSHRDGAKVHLGTRTEKQYETLENQLAF 120  
 DB 61 AENHTHVPFRADDDWSRFLNIGARSFRQALLAYRDGARIHAGTRPGAPQMETADAQLRFL 120  
 QY 121 COOGFSLENALYALSAGVHFTLGCYLEDQEHQVAKEREETPTTDSMP--PILRQAIELFD 178  
 DB 121 CEAFSGADAVNALMTISYFTVGVAVLEEQAGSDSAGER-GGTVEQAPLSPLRAAIDAFD 179  
 QY 179 HOGAEPALFGLGLIICGLEK 200  
 DB 180 EAGPDAFAEQGLAVIVDGLAKR 201

RESULT 7  
 TERS\_ECOLI  
 ID TERS\_ECOLI STANDARD; PRT; 211 AA.  
 AC P21337;  
 DT 01-MAY-1991 (REL. 18, CREATED)  
 DT 01-MAY-1991 (REL. 18, LAST SEQUENCE UPDATE)  
 DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
 DE TETRACYCLINE REPRESSOR PROTEIN CLASS E.  
 GN TETR.  
 OS ESCHERICHIA COLI.  
 OG PLASMID PS1503.  
 OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;  
 OC ENTEROBACTERIACEAE.  
 RN [1]  
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-21.  
 RX MEDLINE; 89201249.  
 RA TOVAR K., ERNST A., HILLEN W.;  
 RA MOL. GEN. GENET. 215:76-80(1988).  
 CC -!- FUNCTION: TETR IS THE REPRESSOR OF THE TETRACYCLINE RESISTANCE  
 CC ELEMENT; ITS AMINO-TERMINAL REGION FORMS A HELIX-TURN-HELIX  
 CC STRUCTURE AND BINDS DNA. BINDING OF TETRACYCLINE TO TETR REDUCES  
 CC THE REPRESSOR AFFINITY FOR THE TETRACYCLINE RESISTANCE GENE  
 CC (TETA) PROMOTER OPERATOR SITES.  
 CC -!- SIMILARITY: BELONGS TO THE TETR/ACRR FAMILY OF TRANSCRIPTIONAL

CC REGULATORS.  
 DR EMBL: M34933; G155020; ALT\_SEQ.  
 DR EMBL: X14035; G43056; -.  
 DR PROSITE: PS01081; HTH\_TETR\_FAMILY; 1.  
 KW TRANSCRIPTION REGULATION; REPRESSOR; DNA-BINDING;  
 KW ANTIBIOTIC RESISTANCE; PLASMID.  
 FT DNA\_BIND 26 45 H-T-H MOTIF (POTENTIAL).  
 FT SITE 64 64 INVOLVED IN BINDING TO [MG-TC]+ (BY  
 FT SIMILARITY).  
 FT METAL 100 100 MAGNESIUM (OF [MG-TC]+ COMPLEX) (BY  
 FT CONFLICT 45 45 SIMILARITY).  
 FT SEQUENCE 211 AA; 23585 MW; F538C813 CRC32;

Query Match 47.2%; Score 509.5; DB 1; Length 211;  
 Best Local Similarity 51.2%; Pred. No. 1.9e-40;  
 Matches 103; Conservative 29; Mismatches 66; Indels 3; Gaps 1;

QY 1 MSRLDKSVINSALELLNEVGIEGLTRKLAOKLGVQPTLYWHVKNRALLDALAIEM 60  
 DB 1 MARLSDDVISMALTLDDSEGLEGLTRKLAOSLAKIEQPTLYWHVKNRQTLNMMLSEAIL 60  
 QY 61 DRHHTHFCPLGESQWDFLRNKAISFRCALLSHRDGAKVHLGTRTEKQYETLENQLAF 120  
 DB 61 AKHTRSAPLPTESWQOFLQENALSFRRKALLVHRDGRALHIGTSPTPPQFEQAQLRCL 120  
 QY 121 COOGFSLENALYALSAGVHFTLGCYLEDQEHQVAKEREETPTTDSMP--PILRQAIELFD 180  
 DB 121 CDAGFSVEALFGLISHSHTLIGAVLEE---OATNQIENNVIDAAPPPLQOEAFFNIQART 177  
 QY 181 GAEPALFGLGLIICGLEK 201  
 DB 178 SAEMAFHFGKSLIFGFSQAL 198

RESULT 8  
 TERS\_VIBAN  
 ID TERS\_VIBAN STANDARD; PRT; 210 AA.  
 AC P51560;  
 DT 01-OCT-1996 (REL. 34, CREATED)  
 DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)  
 DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)  
 DE TETRACYCLINE REPRESSOR PROTEIN CLASS G.  
 GN TETR.  
 OS VIBRIO ANGUILLARUM.  
 OG PLASMID PJA8122.  
 OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;  
 OC VIBRIONACEAE.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-M28122;  
 RX MEDLINE; 93125225.  
 RA ZHAO J., AOKI T.;  
 RL MICROBIOL. IMMUNOL. 36:1051-1060(1992).  
 CC -!- FUNCTION: TETR IS THE REPRESSOR OF THE TETRACYCLINE RESISTANCE  
 CC ELEMENT; ITS AMINO-TERMINAL REGION FORMS A HELIX-TURN-HELIX  
 CC STRUCTURE AND BINDS DNA. BINDING OF TETRACYCLINE TO TETR REDUCES  
 CC THE REPRESSOR AFFINITY FOR THE TETRACYCLINE RESISTANCE GENE  
 CC (TETA) PROMOTER OPERATOR SITES.  
 CC -!- INDUCTION: BY THE [MG-TETRACYCLINE]+ COMPLEX.  
 CC -!- SIMILARITY: BELONGS TO THE TETR/ACRR FAMILY OF TRANSCRIPTIONAL  
 CC REGULATORS.  
 CC EMBL: S52438; G262929; -.  
 DR PROSITE: PS01081; HTH\_TETR\_FAMILY; 1.  
 KW TRANSCRIPTION REGULATION; REPRESSOR; DNA-BINDING;  
 KW ANTIBIOTIC RESISTANCE; PLASMID.  
 FT DNA\_BIND 26 45 H-T-H MOTIF (POTENTIAL).  
 FT SITE 64 64 INVOLVED IN BINDING TO [MG-TC]+ (BY  
 FT SIMILARITY).  
 FT METAL 100 100 MAGNESIUM (OF [MG-TC]+ COMPLEX) (BY  
 FT SEQUENCE 210 AA; 23595 MW; 811CB332 CRC32;

```

DT 01-FEB-1995 (REL. 31, CREATED)
DT 01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)
DE 01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)
GN REGULATORY PROTEIN MTRR.
OS NEISSERIA GONORRHOEA.
OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; AEROBIC RODS AND COCCI;
OC NEISSERIACEAE.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-FA19;
RX MEDLINE; 94254732.
RA PAN W., SPRATT B.G.;
RL MOL. MICROBIOL. 11:769-775(1994).
CC -!- FUNCTION: PUTATIVE REPRESSOR OF MTRC GENE. CONTROLS THE
PERMEABILITY OF THE CELL ENVELOPE TO HYDROPHOBIC COMPOUNDS SUCH AS
ANTIBIOTICS AND DETERGENTS.
CC -!- SIMILARITY: BELONGS TO THE TETR/ACRR FAMILY OF TRANSCRIPTIONAL
REGULATORS.
CC EMBL: Z25797; G438189; -.
DR PIR: S40250; S40250.
DR PROSITE; PS01081; HTH_TETR_FAMILY: 1.
KW TRANSCRIPTION REGULATION; DNA-BINDING; REPRESSOR.
FT DNA_BIND 32 51 H-T-H MOTIF (POTENTIAL).
FT VARIANT 105 105 H -> Y; IN PENICILLIN-RESISTANT ISOLATES.
SQ SEQUENCE 210 AA; 24192 MW; F4CABF98 CRC32;

Query Match 10.2%; Score 110.5; DB 1; Length 210;
Best Local Similarity 27.4%; Pred. No. 0.0021;
Matches 52; Conservative 22; Mismatches 73; Indels 43; Gaps 8;

QY 6 KSKVINSALLELNEVGIEGLTTRKLAQKLGVEQPTLYWHVKNKRALDLALAIEMLDHRHT 65
DB 12 KEHLMLALETFFYKKGARTSLNEIAQAAGVTRGALYWHKNEKEDLDFALFORICDDIEN 71
QY 66 ----HFCPLEGESQWDFLRNKAQKSFRCALLSHRDGAKVHLGTRPTEKQYETLENQALFLC 121
DB 72 CIAQDAADAEGSWTVF-RHTLLHFFERLQSNIDHYKHF-----NILFLKC 116
QY 122 QQGSLENALYALSVAHGFTLGCVLQDQEHQVAKKEERETPTTDSMPPLLRQAIE---LFD 178
DB 117 EH--TEQNA--AVTAIA-----RKHQAIWREKIT-----AVLVEAVENQDLAD 155
QY 179 HOGAEPAFLE 188
DB 156 DLDKETAVIF 165

RESULT 11
UIDR_ECOLI
ID UIDR_ECOLI STANDARD; PRT; 196 AA.
AC Q59431.
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE UID OPERON REPRESSOR (GUS OPERON REPRESSOR).
GN UIDR OR GUSR.
OS ESCHERICHIA COLI.
OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;
OC ENTEROBACTERIACEAE.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-K12;
RA WILSON K.J., JEFFERSON R.A.;
RL SUBMITTED (OCT-1994) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-K12 / MG1655;
RA BLATTNER F.R., PLUNKETT G. III, MAYHEW G.F., PERNA N.T., GLASNER F.D.;
RL SUBMITTED (JAN-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [3]
RP SEQUENCE FROM N.A.

```

```

DT 01-FEB-1995 (REL. 31, CREATED)
DT 01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)
DE 01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)
GN REGULATORY PROTEIN MTRR.
OS NEISSERIA GONORRHOEA.
OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; AEROBIC RODS AND COCCI;
OC NEISSERIACEAE.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-FA19;
RX MEDLINE; 94254732.
RA PAN W., SPRATT B.G.;
RL MOL. MICROBIOL. 11:769-775(1994).
CC -!- FUNCTION: PUTATIVE REPRESSOR OF MTRC GENE. CONTROLS THE
PERMEABILITY OF THE CELL ENVELOPE TO HYDROPHOBIC COMPOUNDS SUCH AS
ANTIBIOTICS AND DETERGENTS.
CC -!- SIMILARITY: BELONGS TO THE TETR/ACRR FAMILY OF TRANSCRIPTIONAL
REGULATORS.
CC EMBL: Z25797; G438189; -.
DR PIR: S40250; S40250.
DR PROSITE; PS01081; HTH_TETR_FAMILY: 1.
KW TRANSCRIPTION REGULATION; DNA-BINDING; REPRESSOR.
FT DNA_BIND 32 51 H-T-H MOTIF (POTENTIAL).
FT VARIANT 105 105 H -> Y; IN PENICILLIN-RESISTANT ISOLATES.
SQ SEQUENCE 210 AA; 24192 MW; F4CABF98 CRC32;

Query Match 10.2%; Score 110.5; DB 1; Length 210;
Best Local Similarity 27.4%; Pred. No. 0.0021;
Matches 52; Conservative 22; Mismatches 73; Indels 43; Gaps 8;

QY 6 KSKVINSALLELNEVGIEGLTTRKLAQKLGVEQPTLYWHVKNKRALDLALAIEMLDHRHT 65
DB 12 KEHLMLALETFFYKKGARTSLNEIAQAAGVTRGALYWHKNEKEDLDFALFORICDDIEN 71
QY 66 ----HFCPLEGESQWDFLRNKAQKSFRCALLSHRDGAKVHLGTRPTEKQYETLENQALFLC 121
DB 72 CIAQDAADAEGSWTVF-RHTLLHFFERLQSNIDHYKHF-----NILFLKC 116
QY 122 QQGSLENALYALSVAHGFTLGCVLQDQEHQVAKKEERETPTTDSMPPLLRQAIE---LFD 178
DB 117 EH--TEQNA--AVTAIA-----RKHQAIWREKIT-----AVLVEAVENQDLAD 155
QY 179 HOGAEPAFLE 188
DB 156 DLDKETAVIF 165

RESULT 11
UIDR_ECOLI
ID UIDR_ECOLI STANDARD; PRT; 196 AA.
AC Q59431.
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE UID OPERON REPRESSOR (GUS OPERON REPRESSOR).
GN UIDR OR GUSR.
OS ESCHERICHIA COLI.
OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;
OC ENTEROBACTERIACEAE.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-K12;
RA WILSON K.J., JEFFERSON R.A.;
RL SUBMITTED (OCT-1994) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-K12 / MG1655;
RA BLATTNER F.R., PLUNKETT G. III, MAYHEW G.F., PERNA N.T., GLASNER F.D.;
RL SUBMITTED (JAN-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [3]
RP SEQUENCE FROM N.A.

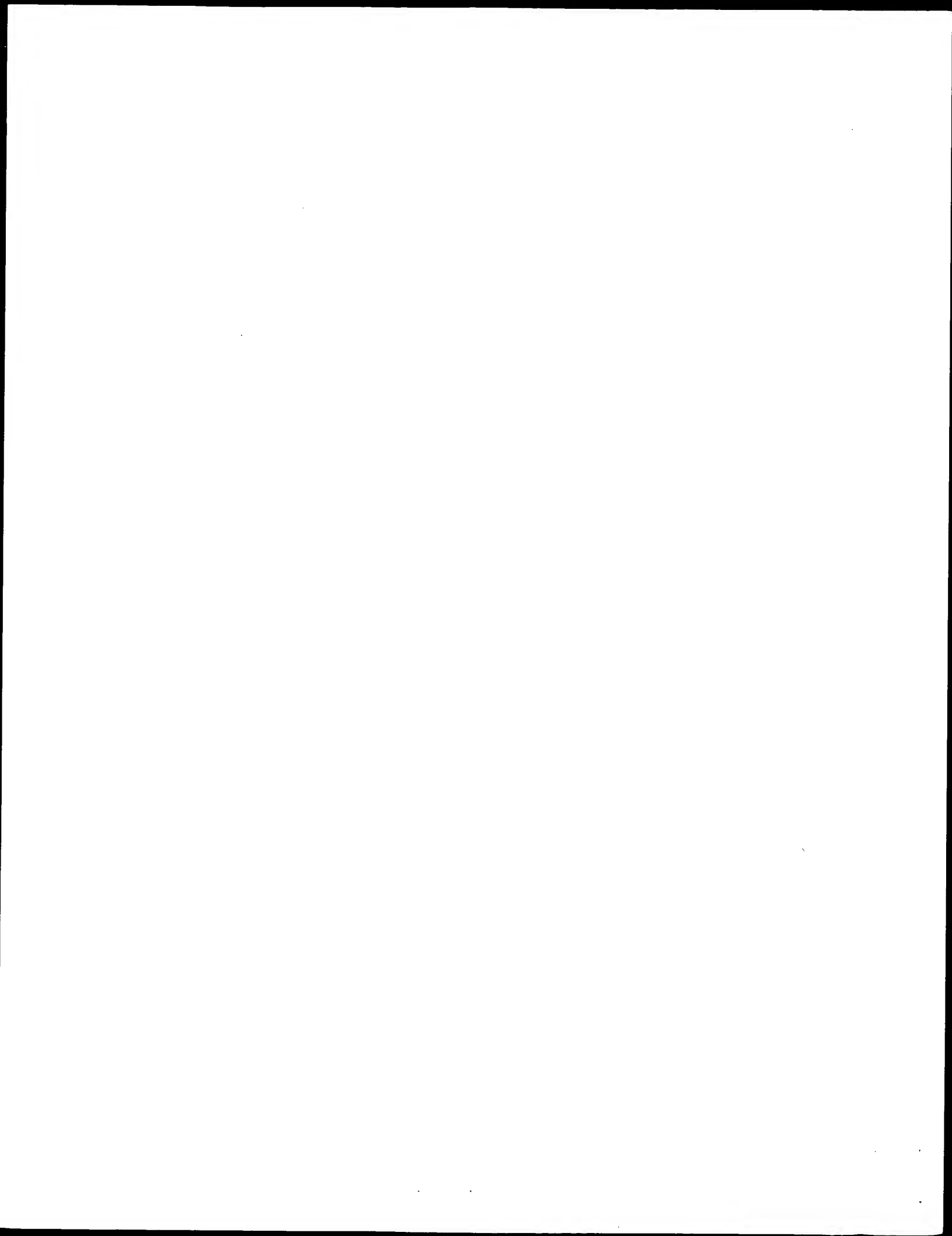
```



```

RESULT 15
BETI_ECOLI
ID BETI_ECOLI STANDARD; PRT; 195 AA.
AC PI7446;
DT 01-AUG-1990 (REL. 15, CREATED)
DT 01-AUG-1990 (REL. 15, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE REGULATORY PROTEIN BETI.
DE BETI.
OS ESCHERICHIA COLI.
OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;
OC ENTEROBACTERIACEAE.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE; 92065800.
RX LAMARK T., KAASEN E., ESHOO M.W., FALKENBERG P., MCDUGALL J.,
RA STROM A.R.;
RA MOL. MICROBIOL. 5:1049-1064(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RC BLATTNER F.R., PLUNKETT G. III, MAYHEW G.F., PERNA N.T., GLASNER F.D.;
RL SUBMITTED (JAN-1997) TO EMBL/GENBANK/DDBJ DATA BANKS.
RN [3]
RP SEQUENCE FROM N.A.
RA DUNCAN M., ALLEN E., ARAUJO R., APARTICIO A.M., CHUNG E., DAVIS K.,
RA FEDERSPIEL N., HYMAN R., KALMAN S., KOMP C., KURDI O., LEW H.,
RA LIN D., NAMATH A., OFENER P., ROBERTS D., SCHRAMM S., DAVIS R.W.;
RL SUBMITTED (OCT-1996) TO EMBL/GENBANK/DDBJ DATA BANKS.
CC -1- FUNCTION: REPRESSOR INVOLVED IN CHOLINE REGULATION OF THE BET
CC GENES.
CC -1- PATHWAY: OSMOREGULATORY CHOLINE-GLYCINE BETAINE PATHWAY.
CC -1- SIMILARITY: BELONGS TO THE TETR/ACRR FAMILY OF TRANSCRIPTIONAL
CC REGULATORS.
CC EMBL; X52905; G48717; -.
CC EMBL; AE000138; G1786505; -.
CC EMBL; U73857; G1657511; ALT_INIT.
CC PIR; S10899; S10899.
CC PIR; S15180; S15180.
CC ECOGENE; EG10111; BETI.
CC PROSITE; PS01081; HTH_TETR_FAMILY; 1.
CC TRANSPOSITION REGULATION: DNA-BINDING; REPRESSOR; TRANS-ACTING FACTOR.

```





GN ACT11.  
 OS STREPTOMYCES COELICOLOR  
 OC BACTERIA; FIRMICUTES; ACTINOBACTERIA; STREPTOMYCETES;  
 RN STREPTOMYCETACEAE; STREPTOMYCETES.  
 RP [1]  
 RX SEQUENCE FROM N.A.  
 RA MEDLINE; 91347376.  
 RA FERNANDEZ-MORENO M.A., CABALLERO J.L., HOPWOOD D.A., MALPARTIDA F.;  
 RT "The act cluster contains regulatory and antibiotic export genes,  
 RT direct targets for translational control by the bldA trna gene of  
 Streptomyces";  
 RL CELL 66:769-780(1991).  
 DR EMBL; M64683; G455360; .  
 DR PFAM; PF00440; tetr; 1.  
 SQ SEQUENCE 259 AA; 28268 MW; 16ECA071 CRC32;

Query Match 13.8%; Score 149.5; DB 9; Length 259;  
 Best Local Similarity 26.0%; Pred. No. 4.5e-06;  
 Matches 57; Conservative 35; Mismatches 90; Indels 37; Gaps 8;  
 QY 4 LDKSKVNSALELLNEVGIEGLTRKLAQKLGVEOPTLYWHVKNKRALD-----AI 57  
 DB 31 LTQDRKIVATAVRVLDAEGDAKFSRLLAELNVTMPSVYVANKDLDLDAVAG 90  
 QY 58 EMLDRHHTHFCPLGESQWDFLRNKAKSFRCALLSHRDGAKVHLGTRPT-EKQYETLENQ 116  
 DB 91 EVPE-----PEPCR-WAEQVKEMCSLRMFLAHRDLARIAIDRVPLGPNMGVMGERT 142  
 QY 117 LAFLCQOQFSLNALYALSVAHGHTLGCVLDEQHQVAKEE-----RET 160  
 DB 143 MNLLRSGHLDELAAAYGDLTLSTVTAELQSSRNPTGEGRQAGVFADQLHGVLKSL 202  
 QY 161 PTDSMPPLRQA--IELFDHQGAEPALFGLGLELIICGL 197  
 DB 203 PAT-SFPNLVHAGPITSLD---SDRRFELGLEIITAGL 237

RESULT 3  
 ID 052558 PRELIMINARY; PRT; 242 AA.  
 AC 052558;  
 DT 01-JUN-1998 (TREMREL. 06, CREATED)  
 DT 01-JUN-1998 (TREMREL. 06, LAST SEQUENCE UPDATE)  
 DT 01-AUG-1998 (TREMREL. 07, LAST ANNOTATION UPDATE)  
 DE REPRESSOR.  
 GN RIFQ.  
 OS AMYCOTOPSIS MEDITERRANEI.  
 OC BACTERIA; FIRMICUTES; ACTINOBACTERIA; PSEUDONOCARDIACEAE;  
 RN AMYCOTOPSIS.  
 RP [1]  
 RX SEQUENCE FROM N.A.  
 RC STRAIN=5699;  
 RA KIM C.G., YU T.W., FRYHLE C., HANDA S., FLOSS H.G.;  
 RL J. BIOL. CHEM. 0:0-0(1998).  
 RP [2]  
 RX SEQUENCE FROM N.A.  
 RC STRAIN=5699;  
 RA AUGUST P.R., TANG L., YOON Y.J., NING S., MUELLER R., HUTCHINSON C.R.,  
 RA TAYLOR M., HOFFMANN D., KIM C.G., ZHANG X., FLOSS H.G.;  
 RL FLOSS H.G.;  
 RA CHEM. BIOL. 5:0-0(0002).  
 RP [3]  
 RX SEQUENCE FROM N.A.  
 RC STRAIN=5699;  
 RA AUGUST P.R., TANG L., YOON Y.J., NING S., MUELLER R., HUTCHINSON C.R.,  
 RA FLOSS H.G.;  
 RL SUBMITTED (DEC-1997) TO EMBL/GENBANK/DBJ DATA BANKS.  
 DR EMBL; AF040570; G2792330; .  
 SQ SEQUENCE 242 AA; 27049 MW; 324EB570 CRC32;

Query Match 12.0%; Score 129.5; DB 9; Length 242;

Best Local Similarity 26.2%; Pred. No. 0.00025;  
 Matches 48; Conservative 35; Mismatches 71; Indels 29; Gaps 7;  
 QY 2 SRLDKSKVNSALELLNEVGIEGLTRKLAQKLGVEOPTLYWHVKNKRAL-----LDALAI 57  
 DB 27 SEDRKRIVATAVRVLDAEGDAKFSRLLAELNVTMPSVYVANKDLDLDAVAG 86  
 QY 58 EM-----LDRHHTHFCPLGESQWDFLRNKAKSFRCALLSHRDGAKV---HLGTRTEQY 110  
 DB 87 EIEPLSLD-----DGDWRDLRALARAWRTTWAHPWAIROYGEILNIGPSLRF 137  
 QY 111 EYLENOLAFLCQOQFSLNALYALSVAHGHTLGCVLDEQ---EHQV---AKEERETPTT 163  
 DB 138 --TECAQAVMACSPLPLKDRSAALNVFQYVYGTATESRWLEHLAETGRTAEFAEAVT 195  
 QY 164 DSM 166  
 DB 196 GSM 198

RESULT 4  
 ID 034892 PRELIMINARY; PRT; 191 AA.  
 AC 034892;  
 DT 01-JAN-1998 (TREMREL. 05, CREATED)  
 DT 01-JAN-1998 (TREMREL. 05, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1998 (TREMREL. 08, LAST ANNOTATION UPDATE)  
 DE TRANSCRIPTION REGULATOR.  
 GN YOB.  
 OS BACILLUS SUBTILIS.  
 OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; BACILLACEAE;  
 RN BACILLUS.  
 RP SEQUENCE FROM N.A.  
 RA LAPIDUS A., GALLERON N., SOROKIN A., EHRLICH D.;  
 RL SUBMITTED (NOV-1997) TO EMBL/GENBANK/DBJ DATA BANKS.  
 RP [2]  
 RX SEQUENCE FROM N.A.  
 RC STRAIN=168;  
 RX MEDLINE; 98044033.  
 RA KUNST F., OGASAWARA N., MOSZER I., ALBERTINI A.M., ALLONI G.,  
 RA AZEVEDO V., BERTERO M.G., BESSIERES P., BOLOTIN A., BORCHERT S.,  
 RA BORRIS R., BOURSIER L., BRANS A., BRAUN M., BRIGNELL S.C., BRON S.,  
 RA BROUILLET S., BRUSCHI C.V., CALDWELL B., CAPUANO V., CARTER N.M.,  
 RA CHOI S.K., CODANI J.J., CONNERTON I.F., CUMMINGS N.J., DANIEL R.A.,  
 RA DENIZOT F., DEVINE K.M., DUSTERHOFT A., EHRLICH S.D., EMERSON P.T.,  
 RA ENTIAN K.D., ERRINGTON J., FABRET C., FERRARI E., FOULGER D., FRITZ C.,  
 RA FUJITA M., FUJITA Y., FUMA S., GALIZZI A., GALLERON N., GHIM S.Y.,  
 RA GLASER P., GOFFEAU A., GOLIGHTLY E.J., GRANDI G., GUISEPPI G.,  
 RA GUY B.J., HAGA K., HATECH J., HARWOOD C.R., HENAUT A., HILBERT H.,  
 RA HOLSAPPEL S., HOSONO S., HULLO M.F., ITAYA M., JONES L., JORIS B.,  
 RA KARAMATA D., KASAHARA Y., KLAER-BLANCHARD M., KLEIN C., KOBAYASHI Y.,  
 RA KOETTER P., KONINGSTEIN G., KROGH S., KUMANO M., KURITA K., LAPIDUS A.,  
 RA LARDINOIS S., LAUBER J., LAZAREVIC V., LEE S.M., LEVINE A., LIU H.,  
 RA MASUDA S., MAUEL C., MEDIGUE C., MEDINA N., MEILLADO R.P., MIZUNO M.,  
 RA MOESTL D., NAKAI S., NOBACK M., NOONE D., O'REILLY M., OGAWA K.,  
 RA OGAWARA A., OUDEGA B., PARK S.H., PARKO V., POHL T.M., PORTETELLE D.,  
 RA PORWOLLIK S., PRESCOTT A.M., PRESECAN E., PUJIC P., PURNELLE B.,  
 RA RAPOLLOK G., REY M., REYNOLDS S., RIEGER M., RIVOLTA C., ROCHA E.,  
 RA ROCHE B., ROSE M., SADAIE Y., SATO T., SCANLAN E., SCHLEICH S.,  
 RA SCHROETER R., SCOFFONE F., SEKIGUCHI J., SEKOWSKA A., SEROR S.J.,  
 RA SERROR P., SHIN B.S., SOLDI B., SOROKIN A., TACCONI E., TAKAGI T.,  
 RA TAKAHASHI H., TAKEMARU K., TAKEUCHI M., TAMAKOSHI A., TANAKA T.,  
 RA TERPSTRA P., TOGNONI A., TOSATO V., UCHIYAMA S., VANDENBOL M.,  
 RA VANNER P., VASSAROTTI A., VIARI A., WAMBUIT R., WEDLER H.,  
 RA WEITENEGER T., WINTERS P., WIPAT A., YAMAMOTO H., YAMANE K.,  
 RA YASUMOTO K., YATA K., YOSHIDA K., YOSHIKAWA H.F., ZUMSTEIN E.,  
 RA YOSHIKAWA H., DANCHIN A.;  
 RT "The complete genome sequence of the gram-positive bacterium Bacillus  
 subtilis";  
 RL NATURE 390:249-256(1997).  
 RP [3]  
 RX SEQUENCE FROM N.A.



178 DHOQREPAFLF 188  
+ + + + +  
Db 155 DDLDKETAVIF 165

RESULT 6  
O34643 PRELIMINARY; PRT; 189 AA.  
ID O34643;  
AC O34643;  
DT 01-JAN-1998 (TREMREL. 05, CREATED)  
DT 01-JAN-1998 (TREMREL. 05, LAST SEQUENCE UPDATE)  
DT 01-NOV-1998 (TREMREL. 08, LAST ANNOTATION UPDATE)  
DYKB.  
DN YVKB.  
GE  
OS BACILLUS SUBTILIS.  
OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; BACILLACEAE;  
OC BACILLUS.  
[1]  
RP SEQUENCE FROM N.A.  
RP LAZAREVIC V., SOLDÓ B., RIVOLTA C., REYNOLDS S., MAUEL C.,  
RA KARAMATA D.;  
RL SUBMITTED (NOV-1997) TO EMBL/GENBANK/DBJ DATA BANKS.  
[2]  
RP SEQUENCE FROM N.A.  
RP STRAIN=168;  
RC MEDLINE; 98044033.  
RX KUNST F., OGASAWARA N., MOSZER I., ALBERTINI A.M., ALLONI G.,  
RA AZEVEDO V., BERTERO M.G., BESSIERES P., BOLOTIN A., BORCHERT S.,  
RA BORRIS R., BOURSTIER L., BRANS A., BRAUN M., BRIGHELL S.C., BRON S.,  
RA BROUILLET S., BRUSCHI C.V., CALDWELL B., CAPUANO V., CARTER N.M.,  
RA CHOI S.K., CODANI J.J., CONNERTON I.F., CUMMINGS N.J., DANIEL R.A.,  
RA DENIZOT F., DEVINE K.M., DUSTERHOFT A., EHRLICH S.D., EMERSON P.T.,  
RA ENTIAN K.D., ERRINGTON J., FABRET C., FERRARI E., FOULGER D., FRITZ C.,  
RA FUJITA M., FUJITA Y., FUMA S., GALIZZI A., GALLERON N., GHIM S.Y.,  
RA GLASER P., GOUFEAU A., GOLIGHTLY E.J., GRANDI G., GUISEPPI G.,  
RA GUY B.J., HAGA K., HALECH J., HARWOOD C.R., HENAUT A., HILBERT H.,  
RA HOLSAPPEL S., HOSONO S., HULLO M.F., ITATA M., JONES L., JORIS B.,  
RA KARAMATA D., KASAHARA Y., KLAER-BLANCHARD M., KLEIN C., KOBAYASHI Y.,  
RA KOTTER P., KONINGSTEIN G., KROCH S., KUNANO M., KURITA K., LAPIDUS A.,  
RA LARDINOIS S., LAUBER J., LAZAREVIC V., LEE S.M., LEVINE A., LIU H.,  
RA MASUDA S., MAUEL C., MEDIGUE C., MEDINA N., MELLADO R.P., MIZUNO M.,  
RA MOREST D., NAKAI S., NOBACK M., NOONE D., O'REILLY M., OGAWA K.,  
RA OGIWARA A., OUDEGA B., PARK S.H., PARKO V., POHL T.M., PORTELLE D.,  
RA ROPWOLLIK S., PRESCOTT A.M., PRESECAN E., PUIC P., PURNELLE B.,  
RA RAPOLLOTT G., REY M., REYNOLDS S., RIEGER M., RIVOLTA C., ROCHA E.,  
RA ROCHE B., ROSE M., SADAIE Y., SATO T., SCANLAN E., SCHLEICH S.,  
RA SCHROETER R., SCOFFONE F., SEKIGUCHI J., SEKOWSKA A., SEROR S.J.,  
RA SERROR P., SHIN B.S., SOLDÓ B., SOROKIN A., TACCONI E., TAKAGI T.,  
RA TAKAHASHI H., TAKEMARU K., TAKEUCHI M., TAMAKOSHI A., TANAKA T.,  
RA TERPSTRA P., TOGNONI A., TOSATO V., UCHIYAMA S., VANDENBOL M.,  
RA VANNIER F., VASSAROTTI A., VIARI A., WAMBUUT R., WEDLER E., WEDLER H.,  
RA WEITZENEGER T., WINTERS P., WIPAT A., YAMAMOTO H., YAMANE K.,  
RA YASUMOTO K., YATA K., YOSHIDA K., YOSHIKAWA H.F., ZUMSTEIN E.,  
RA YOSHIKAWA H., DANCHIN A.;  
RT "The complete genome sequence of the gram-positive bacterium *Bacillus subtilis*,"  
RL NATURE 390:249-256(1997).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=168;  
RA KUNST F., OGASAWARA N., YOSHIKAWA H., DANCHIN A.;  
RL SUBMITTED (NOV-1997) TO EMBL/GENBANK/DBJ DATA BANKS.  
DR EMBL; AF017113; G2618838; "  
DR EMBL; 299122; G1184426; "  
DR PFAM; PF0040; tetr; 1.  
SQ SEQUENCE 189 RA; 21534 MW; 75FBB90 CRC32;

Query Match 10.0%; Score 108.5; DB 9; Length 189;  
Best Local Similarity 19.6%; Pred. No. 0.014;  
Matches 41; Conservative 32; Mismatches 51; Indels 85; Gaps

```

Db 2 RPTNKRILDAAMQLVKKGYRATITKIEAKANVSEATIFRNFKNQGLVEALLSQHSSN 61
QY 63 HHTHFCPLGESWQDFURNKAKSFRCALLSHRDGAKVHLGTRTEKQYETLENQALFLCQ 122
Db 62 RGSILEQTEGDKYKDL-----HIGT----- 83
QY 123 QGFSLENALVALSAVGHFTLGCVCLEDOHQ-----VAKEERETPTDSM 166
Db 83 -----CLLELEHRKDIKISFREPAFMQDQVINHVTEYP-QSM 119
QY 167 PPLLRQAIELFDHQ-----AEPALF 188
Db 120 KOLLVDYLTGMGKGVQIOTGNEAHDV 148

RESULT 7
O67927 PRELIMINARY; PRT; 192 AA.
AC O67927;
DT 01-AUG-1998 (TREMBREL. 07, CREATED)
DT 01-AUG-1998 (TREMBREL. 07, LAST SEQUENCE UPDATE)
DE 01-NOV-1998 (TREMBREL. 08, LAST ANNOTATION UPDATE)
GN ACRR2.
OS AQUIFEX AEOLICUS.
OC BACTERIA; AQUIFICALES; AQUIFICACEAE; AQUIFEX.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VF5;
RX MEDLINE; 98196666.
RA DECKERT G., WARREN P.V., GAASTERLAND T., YOUNG W.G., LENOX A.L.,
RA GRAHAM D.E., OVERBEK R., SNEAD M.A., KELLER M., AUJAY M., HUBER R.,
RA FELDMAN R.A., SHORT J.M., OLSON G.J., SWANSON R.V.;
RL SUBMITTED (JUL-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL; AE000776; G2984362;
DR PROSITE; PS01081; HTH_TETR_FAMILY; 1.
SQ SEQUENCE 192 AA; 22259 MW; D2E83A94 CRC32;

Query Match 9.4%; Score 101.5; DB 9; Length 192;
Best Local Similarity 22.7%; Pred. No. 0.06;
Matches 45; Conservative 34; Mismatches 68; Indels 51; Gaps 8;

QY 6 KSKVINSALLENEVGIEGLTRKLAQKLGVEOPTLYWHVKNK-----RALLDALAIEMLD 61
Db 15 KEILSALKLFSKKGFKETIKINDIAKEVGITEGAIYRHTFSKEEIIKSLSEITKEL-- 73
QY 62 RHHTFCPLGESWQDFLR-----NKAKSFRCALLSHRDGAKVHLGTRTEKQY 110
Db 73 RHKLEVALQGTDEEILEISIVDTLIDYAFSPESFRFLNLYH-----LLKEY 120
QY 111 ETLENO-----LAFLCQGFLENALY-----ALSAVGHFTLGCVCLEDOHQVAKER----- 159
Db 121 GEVKNLPGLIILKFL-----NGLYLRKRLKTYPEITALAVTGSVERVIFKERNELD 172
QY 159 --ETPTDSMPPLLRQAI 174
Db 173 YDEETIKKELKVKLSAI 190

RESULT 8
O53789 PRELIMINARY; PRT; 196 AA.
ID O53789

```

```

AC O53789;
DT 01-JUN-1998 (TREMBREL. 06, CREATED)
DT 01-JUN-1998 (TREMBREL. 06, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBREL. 08, LAST ANNOTATION UPDATE)
DE PUTATIVE TRANSCRIPTIONAL REGULATOR.
GN MV040.09.
OS MYCOBACTERIUM TUBERCULOSIS.
OC BACTERIA; FIRMICUTES; ACTINOBACTERIA; MYCOBACTERIA; MYCOBACTERIACEAE;
OC MYCOBACTERIUM.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RA HAMLIN N., CHURCHER C.M.;
RL SUBMITTED (FEB-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RA COLE S.T., PARKHILL J., BARRELL B.G., RAJANDREAM M.A.;
RL SUBMITTED (FEB-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RX MEDLINE; 96181548.
RA PHILIPP W.J., POULET S., EIGLMEIER K., PASCOPELLA L.,
RA BALASUBRAMANIAN V., HEYM B., BERGH S., BLOOM B.R., JACOBS W.R. JR.,
RA COLE S.T.;
RT "An integrated map of the genome of the tubercle bacillus,
RT Mycobacterium tuberculosis H37Rv, and comparison with Mycobacterium
RT leprae.";
RL PROC. NATL. ACAD. SCI. U.S.A. 93:3132-3137(1996).
DR EMBL; AL021943; E1253107;
SQ SEQUENCE 196 AA; 21262 MW; 17D86001 CRC32;

Query Match 9.1%; Score 98; DB 9; Length 196;
Best Local Similarity 27.3%; Pred. No. 0.13;
Matches 42; Conservative 24; Mismatches 74; Indels 14; Gaps 5;

QY 2 SRLDKSKVINSALLENEVGIEGLTRKLAQKLGVEOPTLYWHVKNKRALDALAIEMLD 61
Db 5 AKLSRESIEGALFTLDREGWDSLINALATQLTGKPSLYNHVDSLEDLRRVIRVID 64
QY 62 RHHTFCPL-EGESWQDFLRNKAKSFRCALLSHRDG-----AKVHLGTRTEKQYET--- 113
Db 65 DIIIMLNVRGAGRADDVAVLMAGAYR-STAHHPGYSAFTRMPLGGDDPEYTAATRGA 123
QY 113 LENQALFLCQGFLENALYAL----SAVGHFTL 142
Db 124 AAPVIAVLSYGLDGEQAFYAALFEFWSALHGFLV 157

RESULT 9
Q51597 PRELIMINARY; PRT; 186 AA.
ID Q51597
AC Q51597;
DT 01-NOV-1996 (TREMBREL. 01, CREATED)
DT 01-NOV-1996 (TREMBREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBREL. 08, LAST ANNOTATION UPDATE)
DE CAM REPRESSOR.
GN CAMR.
OS PSEUDOMONAS PUTIDA.
OC PLASMID CAM.
OC BACTERIA; PROTEOBACTERIA; GAMMA SUBDIVISION; PSEUDOMONAS GROUP;
OC PSEUDOMONAS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PPG1;
RX MEDLINE; 86223770.
RA KOGA H., ARAMAKI H., YAMAGUCHI E., TAKEUCHI K., HORIUCHI T.,
RA GUNSAIUS I.C.;
RT "camR, a negative regulator locus of the cytochrome P-450cam
RT hydroxylase operon.";
RL J. BACTERIOL. 166:1089-1095(1986).

```

```

RN  [2]
RP  SEQUENCE FROM N.A.
RX  STRAIN=PPG1;
RC  MEDLINE; 93326643.
RA  ARAMAKI H., KOGA H., SAGARA Y., HOSOI M., HORIUCHI T.;
RT  "Complete nucleotide sequence of the 5'-exo-hydroxycamphor
RT  dehydrogenase gene on the CAM plasmid of Pseudomonas putida (ATCC
RT  17453).";
RL  BIOCHIM. BIOPHYS. ACTA 1174:91-94(1993).
RN  [3]
RP  SEQUENCE FROM N.A.
RX  STRAIN=PPG1;
RC  MEDLINE; 94304997.
RA  ARAMAKI H., SAGARA Y., TAKEUCHI K., KOGA H., HORIUCHI T.;
RT  "Nucleotide sequence of the gene encoding a repressor for the
RT  cytochrome P-450cam hydroxylase operon on the Pseudomonas putida CAM
RT  plasmid.";
RL  BIOCHIMIE 76:63-70(1994).
DR  EMBL; D14680; G473745; -.
DR  PFAM; PF00440; tetr; 1.
KW  PLASMID.
SQ  SEQUENCE 186 AA; 20438 MW; 56E43196 CRC32;

Query Match      8.8%; Score 94.5; DB 9; Length 186;
Best Local Similarity 28.6%; Pred. No. 0.24;
Matches 24; Conservative 20; Mismatches 31; Indels 9; Gaps 3;

QY  6 KSKVNSALELNEVGIEGLTTRKLAQKLGVEOPTLYWHVKNKRALDLAIAEMLD 61
DB  4 KQSLHAAMRLLSAKRGDCAATMRPICAEGVTPPTLYHHYDGLGHLKAAIDETVQVAE 63

QY  62 RH--THFCPLEG--ESWQDLR 80
DB  64 AYHGTEERGPLKGRDGNWATFLQ 87

RESULT 10
P73186 PRELIMINARY; PRT; 236 AA.
AC P73186;
DT 01-FEB-1997 (TREMBLREL. 02, CREATED)
DT 01-FEB-1997 (TREMBLREL. 02, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL 26.4 KD PROTEIN.
OS SYNECHOCYSTIS SP. (STRAIN PCC 6803).
OC BACTERIA; CYANOBACTERIA; CHROCOCCALES; SYNECHOCYSTIS.
RN [1]
RP SEQUENCE FROM N.A.
RA TABATA S.;
RL SUBMITTED (JUN-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE; 97061201.
RA KANEKO T., SATO S., KOTANI H., TANAKA A., ASAMIZU E., NAKAMURA Y.,
RA MIYAJIMA N., HIROSAWA M., SUGIURA M., SASAMOTO S., KIMURA T.,
RA HOSOUCHI T., MATSUNO A., MURAKI A., NAKAZAKI N., NARUO K., OKUMURA S.,
RA SHIMPO S., TAKEUCHI C., WADA T., WATANABE A., YAMADA M., YASUDA M.,
RA TABATA S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synechocystis sp. strain PCC6803. II. Sequence determination of the
RT entire genome and assignment of potential protein-coding regions.";
RL DNA RES. 3:109-136(1996).
DR EMBL; D90904; G1652289; -.
DR PROSITE; PS01081; HTH_TETR_FAMILY; 1.
DR PFAM; PF00440; tetr; 1.
KW HYPOTHETICAL PROTEIN.
SQ SEQUENCE 236 AA; 26384 MW; A8D84E37 CRC32;

Query Match      8.8%; Score 94.5; DB 9; Length 236;
Best Local Similarity 22.5%; Pred. No. 0.33;
Matches 36; Conservative 24; Mismatches 55; Indels 45; Gaps 5;

```

```

QY  8 KVNSALELNEVGIEGLTTRKLAQKLGVEOPTLYWHVKNKRALDLAIAEMLD 67
DB  46 RILKGALKLFGTGGEGTTTKDLAANVAEGTLFRYFTNKKAILEVAT----- 96
QY  68 CPLEGESWQDFLRNKAISFRCALLSHROGAKVHLGTRPTEK-OYETLENQALFLCQOQS 126
DB  96 -----AGWVEILLTD-----LLTELSEMGSYKAIQAQVMK---RRMF 128
QY  127 LENALYALSAVGHFTLGCVLEDEQEHQVAKEREETPTTDSM 166
DB  129 LRENKYLQV-----CFVEAQYHPDELKIQSEIIDKM 161

RESULT 11
O67157 PRELIMINARY; PRT; 179 AA.
ID O67157;
DT 01-AUG-1998 (TREMBLREL. 07, CREATED)
DT 01-AUG-1998 (TREMBLREL. 07, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE TRANSCRIPTIONAL REGULATOR (TETR/ACRR FAMILY).
GN ACRR1.
OS AQUIFEX AEOLICUS.
OC BACTERIA; AQUIFICALES; AQUIFICACEAE; AQUIFEX.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VF5;
RX MEDLINE; 98196666.
RA DECKERT G., WARREN P.V., GAASTERLAND T., YOUNG W.G., LENOX A.L.,
RA GRAHAM D.E., OVERBECK R., SNEAD M.A., KELLER M., AUJAY M., HUBER R.,
RA FELDMAN R.A., SHORT J.M., OLSON G.J., SWANSON R.V.;
RL SUBMITTED (JUL-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL; AE00721; G2983549; -.
SQ SEQUENCE 179 AA; 21554 MW; E359173B CRC32;

Query Match      8.7%; Score 94; DB 9; Length 179;
Best Local Similarity 21.4%; Pred. No. 0.26;
Matches 47; Conservative 33; Mismatches 68; Indels 72; Gaps 9;

QY  6 KSKVNSALELNEVGIEGLTTRKLAQKLGVEOPTLYWHVKNKRALDLAIAEMLD 65
DB  4 KERILEVSKELFFKEGYQGTSVVEIVKRNLSKGAFFHFKSKKEELI---TEIERTHK 59
QY  66 HFCPLEGESWQDFLRNKAISFRCALLSHROGAKVHLGTRPTEKQYETL---ENQALF---- 120
DB  60 KIISL-----FEENKEKTEPELL-----EMPLEVLYREKKVYIFLF 96
QY  120 --LCOQGF-----SLENALYAL-----SAVGHTLGCVLEDEQEHQVAKERE 159
DB  97 DLLCEKFRNTYFEKIEDAKRRFEKFEKHPPSKAEILSEILFLGLRQLILHYVIKERE 156
QY  160 TPTTDSMPPLRQATLFDHOGAEPAPFLGLGLIICGLEK 199
DB  157 LP---FLAEKLE-----GLKLIFEGVKK 177

RESULT 12
O87854 PRELIMINARY; PRT; 217 AA.
ID O87854;
AC O87854;
DT 01-NOV-1998 (TREMBLREL. 08, CREATED)
DT 01-NOV-1998 (TREMBLREL. 08, LAST SEQUENCE UPDATE)

```

[illegible]

QY 65 THFCPL 71  
 Db 69 ----PLE 71

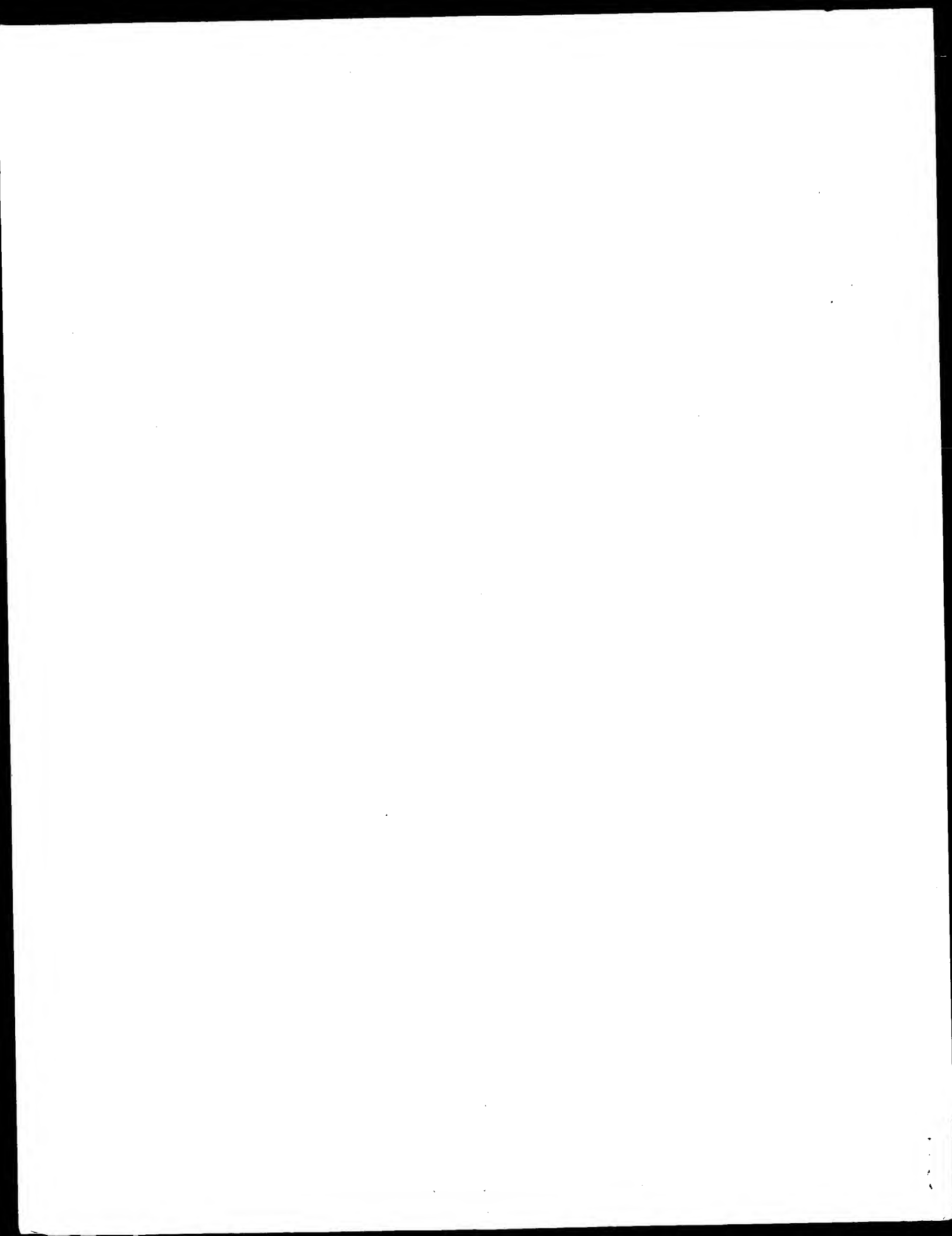
8 KVNSALELLENGVIEGLTTRKLAQKGVQPTLYWHVKNKRALDLALAIEMLDHRHT 65  
 11 KILQAAIEVISEKGLDKASISDIVKKAGTQGTFFLYFSSKNALIPATAENLLT--HT 66

## RESULT 15

O07001  
 AC O07001 PRELIMINARY; PRT; 194 AA.  
 DT 01-JUL-1997 (TREMBLREL. 04, CREATED)  
 DT 01-JUL-1997 (TREMBLREL. 04, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)  
 DE HYPOTHETICAL 22.2 KD PROTEIN.  
 GN YVDT.  
 OS BACILLUS SUTLILIS.  
 OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; BACILLACEAE;  
 CC BACILLUS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA DENIZOT F.C.;  
 PL SUBMITTED (APR-1997) TO EMBL/GENBANK/DBJ DATA BANKS.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-168;  
 RX MEDLINE; 98044033.  
 RA KUNST F., OGASAWARA N., MOSZER I., ALBERTINI A.M., ALLONI G.,  
 RA AZEVEDO V., BERTERO M.G., BESSIERES P., BOLOTIN A., BORCHERT S.,  
 RA BORRIS R., BOURSIER L., BRANS A., BRAUN M., BRIGNELL S.C., BRON S.,  
 RA BROUILLET S., BRUSCHI C.V., CALDWELL B., CAPUANO V., CARTER N.M.,  
 RA CHOI S.K., CODANI J.J., CONNERTON I.F., CUMINGS N.J., DANIEL R.A.,  
 RA DENIZOT F., DEVINE K.M., DUSTERHOFT A., EHRLICH S.D., EMERSON P.T.,  
 RA ENTIAN K.D., ERRINGTON J., FABRET C., FERRARI E., FOULGER D., FRITZ C.,  
 RA FUJITA M., FUJITA Y., FUMA S., GALIZZI A., GALLERON N., GHIM S.Y.,  
 RA GLASER P., GOFFEAU A., GOLIGHTLY E.J., GRANDI G., GUISEPPi G.,  
 RA GUY B.J., HAGA K., HAIECH J., HARWOOD C.R., HENAUT A., HILBERT H.,  
 RA HOLSAPPEL S., HOSONO S., HULLO M.F., ITAYA M., JONES L., JORIS B.,  
 RA KARAYATA D., KASHARA Y., KLAER-BLANCHARD M., KLEIN C., KOBAYASHI Y.,  
 RA KOETTER P., KONINGSTEIN G., KROGH S., KUMANO M., KURITA K., LAPIDUS A.,  
 RA LARDINOIS S., LAUBER J., LAZAREVIC V., LEE S.M., LEVINE A., LIU H.,  
 RA MASUDA S., MAUEL C., MEDIGUE C., MEDINA N., MELLADO R.P., MIZONO M.,  
 RA MESTL D., NAKAI S., NOBACK M., NOONE D., O'REILLY M., OGAWA K.,  
 RA OGIIWARA A., OUDEGA B., PARK S.H., PARRO V., POHL T.N., PORTETELLE D.,  
 RA PORWOLLIK S., PRESCOTT A.M., PRESECAN E., PUJIC P., PURNELLE B.,  
 RA ROCHE B., ROSE M., REYNOLDS S., RIEGER M., RIVOLTA C., ROCHA E.,  
 RA SCHROETER R., SCOFFONE F., SEKIGUCHI J., SEKOWSKA A., SEROR S.J.,  
 RA SERROR P., SHIN B.S., SOLDI B., SOROKIN A., TACCONI E., TAKAGI T.,  
 RA TAKAHASHI H., TAKEMARU K., TAKEUCHI M., TAMAKOSHI A., TANAKA T.,  
 RA TERPSTRA P., TOGNONI A., TOSATO V., UCHIYAMA S., VANDENBOL M.,  
 RA VANNIER F., VASSAROTTI A., VIARI A., WAMBUUT R., WEDLER H.,  
 RA WEITZENEGGER T., WINTERS P., WIPAT A., YAMAMOTO H., YAMANE K.,  
 RA YASUMOTO K., YATA K., YOSHIDA K., YOSHIKAWA H.F., ZUMSTEIN E.,  
 RA YOSHIKAWA H., DANCHIN A.;  
 RT "The complete genome sequence of the gram-positive bacterium Bacillus  
 subtilis."  
 RL NATURE 390:249-256(1997).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-168;  
 RA KUNST F., OGASAWARA N., YOSHIKAWA H., DANCHIN A.;  
 RL SUBMITTED (NOV-1997) TO EMBL/GENBANK/DBJ DATA BANKS.  
 DR EMBL; 294043; E313126;  
 DR EMBL; 295121; E1186136;  
 DR PFAM; PF00440; tetr; 1.  
 KW HYPOTHETICAL PROTEIN.  
 SQ SEQUENCE 194 AA; 22227 MW; D3457AEC CRC32;

Query Match 8.2%; Score 89; DB 9; Length 194;  
 Best Local Similarity 31.0%; Pred. No. 0.8;  
 Matches 18; Conservative 17; Mismatches 21; Indels 2; Gaps 1;

Search completed: June 9, 1999, 13:00:33  
 Job time: 9152 sec



GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 9, 1999, 12:59:17 ; Search time 23.43 Seconds  
(without alignments)  
178.687 Million cell updates/sec

Title: US-08-486-814-19  
Perfect score: 1080  
Sequence: 1 MSRLDKSVINSALELLNEV.....FGLELLIGLEKOLKCESGS 207

Scoring table: BLOSUM62

Searched: 162890 seqs, 20225328 residues

Database : A\_Geneseq\_34:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	1080	100.0	335	1 R89761	Mutated Tn-10 deri
2	1080	100.0	207	1 R89764	Mutated reverse re
3	1080	100.0	335	1 W08293	Mutated Tet repres
4	1080	100.0	207	1 W08297	Protein encoded by
5	1080	100.0	335	1 W71323	Protein encoded by
6	1080	100.0	207	1 W71327	tTA transactivator
7	1058	98.0	335	1 R64808	Wild type Tn-10 de
8	1058	98.0	207	1 R89763	Wild-type E.coli T
9	1058	98.0	207	1 W08296	Multi-chimeric tra
10	1058	98.0	651	1 W48357	Multi-chimeric tra
11	1058	98.0	651	1 W47583	Protein encoded by
12	1058	98.0	207	1 W71326	Sequence of the te
13	1053	97.5	207	1 R48630	tTAS transactivato
14	1053	97.5	297	1 R64809	tTAS transactivato
15	1053	97.5	297	1 R85324	Plasmid PASK75 ope
16	1053	97.5	1088	1 R88636	Tetracycline trans
17	1053	97.5	354	1 W21994	Pleckstrin homolog
18	1053	97.5	349	1 W54311	tTA transactivator
19	1051	97.3	335	1 R85323	pTet-tTAK encoded
20	1048	97.0	336	1 W08323	Mutant tetracyclin
21	1046	96.9	207	1 W08323	Mutated Tet repres
22	1045.5	96.8	206	1 W12401	Mutated Tet repres
23	1045	96.8	207	1 W08317	Mutated Tet repres
24	1045	96.8	207	1 W08318	Mutated Tet repres
25	1045	96.8	207	1 W08300	Mutated Tet repres
26	1044	96.7	207	1 W08320	Mutated Tet repres
27	1044	96.7	207	1 W08304	Mutated Tet repres
28	1043	96.6	207	1 W08312	Mutated Tet repres
29	1041	96.4	207	1 W08306	Mutated Tet repres
30	1039	96.2	207	1 W08325	Mutated Tet repres
31	1039	96.2	207	1 W08326	Mutated Tet repres
32	1037	96.0	207	1 W08301	Mutated Tet repres
33	1036	95.9	207	1 W08321	Mutated Tet repres
34	1036	95.9	207	1 W08305	Mutated Tet repres
35	1035	95.8	207	1 W08303	Mutated Tet repres
36	1035	95.8	207	1 W08310	Mutated Tet repres
37	1035	95.8	207	1 W08313	Mutated Tet repres
38	1034	95.7	207	1 W08316	Mutated Tet repres
39	1034	95.7	207	1 W08319	Mutated Tet repres
40	1034	95.7	207	1 W08302	Mutated Tet repres
41	1034	95.7	207	1 W08308	Mutated Tet repres
42	1033	95.6	207	1 W08322	Mutated Tet repres
43	1033	95.6	207	1 W08324	Mutated Tet repres

## ALIGNMENTS

RESULT	ID	Query Match	Score	Length	DB	ID	Description
1	R89761	100.0%	100.0%	335	AA	R89761	Mutated Tn-10 deri
2	R89761	100.0%	100.0%	335	AA	R89761	Mutated reverse re
3	R89761	100.0%	100.0%	335	AA	R89761	Mutated Tet repres
4	R89761	100.0%	100.0%	335	AA	R89761	Protein encoded by
5	R89761	100.0%	100.0%	335	AA	R89761	Protein encoded by
6	R89761	100.0%	100.0%	335	AA	R89761	tTA transactivator
7	R89761	100.0%	100.0%	335	AA	R89761	Wild type Tn-10 de
8	R89761	100.0%	100.0%	335	AA	R89761	Wild-type E.coli T
9	R89761	100.0%	100.0%	335	AA	R89761	Multi-chimeric tra
10	R89761	100.0%	100.0%	335	AA	R89761	Multi-chimeric tra
11	R89761	100.0%	100.0%	335	AA	R89761	Protein encoded by
12	R89761	100.0%	100.0%	335	AA	R89761	Sequence of the te
13	R89761	100.0%	100.0%	335	AA	R89761	tTAS transactivato
14	R89761	100.0%	100.0%	335	AA	R89761	tTAS transactivato
15	R89761	100.0%	100.0%	335	AA	R89761	Plasmid PASK75 ope
16	R89761	100.0%	100.0%	335	AA	R89761	Tetracycline trans
17	R89761	100.0%	100.0%	335	AA	R89761	Pleckstrin homolog
18	R89761	100.0%	100.0%	335	AA	R89761	tTA transactivator
19	R89761	100.0%	100.0%	335	AA	R89761	pTet-tTAK encoded
20	R89761	100.0%	100.0%	335	AA	R89761	Mutant tetracyclin
21	R89761	100.0%	100.0%	335	AA	R89761	Mutated Tet repres
22	R89761	100.0%	100.0%	335	AA	R89761	Mutated Tet repres
23	R89761	100.0%	100.0%	335	AA	R89761	Mutated Tet repres
24	R89761	100.0%	100.0%	335	AA	R89761	Mutated Tet repres
25	R89761	100.0%	100.0%	335	AA	R89761	Mutated Tet repres
26	R89761	100.0%	100.0%	335	AA	R89761	Mutated Tet repres
27	R89761	100.0%	100.0%	335	AA	R89761	Mutated Tet repres
28	R89761	100.0%	100.0%	335	AA	R89761	Mutated Tet repres
29	R89761	100.0%	100.0%	335	AA	R89761	Mutated Tet repres
30	R89761	100.0%	100.0%	335	AA	R89761	Mutated Tet repres
31	R89761	100.0%	100.0%	335	AA	R89761	Mutated Tet repres
32	R89761	100.0%	100.0%	335	AA	R89761	Mutated Tet repres
33	R89761	100.0%	100.0%	335	AA	R89761	Mutated Tet repres
34	R89761	100.0%	100.0%	335	AA	R89761	Mutated Tet repres
35	R89761	100.0%	100.0%	335	AA	R89761	Mutated Tet repres
36	R89761	100.0%	100.0%	335	AA	R89761	Mutated Tet repres
37	R89761	100.0%	100.0%	335	AA	R89761	Mutated Tet repres
38	R89761	100.0%	100.0%	335	AA	R89761	Mutated Tet repres
39	R89761	100.0%	100.0%	335	AA	R89761	Mutated Tet repres
40	R89761	100.0%	100.0%	335	AA	R89761	Mutated Tet repres
41	R89761	100.0%	100.0%	335	AA	R89761	Mutated Tet repres
42	R89761	100.0%	100.0%	335	AA	R89761	Mutated Tet repres
43	R89761	100.0%	100.0%	335	AA	R89761	Mutated Tet repres

DT 07-JUL-1996 (first entry)  
 DE Mutated Tn-10 derived Tet repressor.  
 KW Tet repressor; Herpes simplex virus; HSV; virion protein 16; VP16;  
 KW fusion protein; gene expression; regulation; inhibition; activation;  
 OS transcription.  
 PN WO9601313-A1.  
 PD 18-JAN-1996.  
 PF 29-JUN-1995; U08179.  
 PR 01-JUL-1994; US-270637.  
 PR 15-JUL-1994; US-275876.  
 PR 03-FEB-1995; US-383754.  
 PR 07-JUN-1995; US-486814.  
 PA (BUJA)/ BUJARD H.  
 PA (GOS)/ GOSSEN M.  
 PI Bujard H, Gossen M;  
 DR WPI: 96-087666/09.  
 DR N-PSDB; T11352.  
 DT New tetracycline-regulated transcription modulators - comprising  
 PT fusion proteins which bind to tet operator sequences to activate or  
 PT inhibit transcription  
 PS Claim 44: Page 80; 112pp; English.  
 CC Fusion proteins comprising a first polypeptide which binds to a tet  
 CC operator sequence in the presence of tetracycline or a tetracycline  
 CC analogue, operatively linked to a second polypeptide which either  
 CC activates or inhibits transcription in eukaryotic cells may be used  
 CC to activate or inhibit transcription. Such proteins may be used to  
 CC regulate gene expression in cells and may be particularly useful for  
 CC gene therapy and for expression of gene products in transgenic  
 CC organisms. Induction of gene expression is rapid, efficient and  
 CC strong, typically 1000-2000 fold. The inducing agent does not cause  
 CC pleiotropic effects or cytotoxicity in eukaryotic cells. This  
 CC sequence is a mutated Tn-10 tet repressor used in a transcription  
 CC inhibiting fusion protein.  
 SQ Sequence 207 AA;

Query Match 100.0%; Score 1080; DB 1; Length 207;  
 Best Local Similarity 100.0%; Pred. No. 3.4e-115;  
 Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MSRLDKSVINSALLELNEVGIEGLTTRKLAQKLGVEQPTLYWHVKNRALLDALATEML 60  
 DB 1 MSRLDKSVINSALLELNEVGIEGLTTRKLAQKLGVEQPTLYWHVKNRALLDALATEML 60  
 QY 61 DRHHTHFCPLKGSQWDFLRNKAQKSHRNGAKVHSDTRPTEKQYETLENQALFL 120  
 DB 61 DRHHTHFCPLKGSQWDFLRNKAQKSHRNGAKVHSDTRPTEKQYETLENQALFL 120  
 QY 121 CQOGFSLNLYALSAYGHFTLGCVLDEQEHQVAKERETPTTDSMPPLLRQAIELFDHQ 180  
 DB 121 CQOGFSLNLYALSAYGHFTLGCVLDEQEHQVAKERETPTTDSMPPLLRQAIELFDHQ 180  
 QY 181 GAEPAPFLGLELIICGLEKQKCESGS 207  
 DB 181 GAEPAPFLGLELIICGLEKQKCESGS 207

RESULT 3  
 W08293  
 ID W08293 standard; Protein; 335 AA.  
 AC W08293;  
 DT 18-MAR-1997 (first entry)  
 DE Mutated reverse tet repressor/VP16 fusion protein.  
 KW Tet; tetracycline; Tc; operator; transcription; regulation; inducible;  
 KW repressor; gene expression; therapy; transgenic animal; disease model.  
 OS Synthetic.  
 PN Key  
 DR Location/Qualifiers  
 domain 1..207  
 FT /note= "reverse tet repressor"  
 FT /label= substitution  
 FT /note= "Iys replaces wild-type Glu"

FT misc\_difference 95  
 FT /label= substitution  
 FT /note= "Asn replaces wild-type Asp"  
 FT misc\_difference 101  
 FT /label= substitution  
 FT /note= "Ser replaces wild-type Leu"  
 FT misc\_difference 102  
 FT /label= substitution  
 FT /note= "Asp replaces wild-type Gly"  
 FT 208..335  
 FT /note= "C-terminal fragment of VP16 protein"

PN WO9640892-A1.  
 PD 19-DEC-1996.  
 PF 06-JUN-1996; U09049.  
 PR 07-JUN-1995; US-485971.  
 PA (BADI) BASF AG.  
 PI Bujard H, Gossen M, Helbl V, Hallen W, Schnappinger D;  
 DR WPI: 97-052305/05.  
 DT Nucleic acid encoding tetracycline-inducible transcription  
 PT regulatory fusion protein - comprising modified tetracycline  
 PT repressor able to bind mutant tet operator, fused to transcription  
 PT regulator, useful for modulating eukaryotic gene expression  
 PS Example 1: Page 77-78; 117pp; English.  
 CC W08293 represents a "reverse" Tet repressor (rtetr), i.e. a wild-type  
 CC TetR protein mutated so as to bind its target in the presence  
 CC rather than the absence of tetracycline (TC), fused to a C-terminal  
 CC polypeptide of VP16 (herpes simplex virus virion protein 16), a  
 CC transcriptional activator. The fusion protein forms a Tc-controlled  
 CC specificational activator (tTA). The main invention of the  
 CC specification concerns modified TetR proteins that bind to modified  
 CC class B tet operator sequences tetO-4C and tetO-6C (see T45711 and  
 CC T48478). Modified TetR proteins can be fused to any transcription  
 CC regulatory polypeptide and used to control transcription of a tetO-4C  
 CC or tetO-6C linked gene. Nucleic acid encoding such a fusion protein  
 CC may be introduced into a cell and transcription of the protein can be  
 CC controlled by altering the concn. of tetracycline (or an analogue) in  
 CC the cell, as appropriate. This ability to modulate gene expression in  
 CC a predictable way is very useful in gene therapy and for recombinant  
 CC protein prodn. in cultured cells or transgenic animals. The  
 CC Tc-inducible system is also useful for the prodn. of transgenic animal  
 CC models for the study of disease and also for the study of gene  
 CC function e.g. during differentiation. The Tc-inducible system allows  
 CC rapid activation of gene transcription without cellular toxicity, high  
 CC concns. of inducer are not required.  
 SQ Sequence 335 AA;

Query Match 100.0%; Score 1080; DB 1; Length 335;  
 Best Local Similarity 100.0%; Pred. No. 6.8e-115;  
 Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MSRLDKSVINSALLELNEVGIEGLTTRKLAQKLGVEQPTLYWHVKNRALLDALATEML 60  
 DB 1 MSRLDKSVINSALLELNEVGIEGLTTRKLAQKLGVEQPTLYWHVKNRALLDALATEML 60  
 QY 61 DRHHTHFCPLKGSQWDFLRNKAQKSHRNGAKVHSDTRPTEKQYETLENQALFL 120  
 DB 61 DRHHTHFCPLKGSQWDFLRNKAQKSHRNGAKVHSDTRPTEKQYETLENQALFL 120  
 QY 121 CQOGFSLNLYALSAYGHFTLGCVLDEQEHQVAKERETPTTDSMPPLLRQAIELFDHQ 180  
 DB 121 CQOGFSLNLYALSAYGHFTLGCVLDEQEHQVAKERETPTTDSMPPLLRQAIELFDHQ 180  
 QY 181 GAEPAPFLGLELIICGLEKQKCESGS 207  
 DB 181 GAEPAPFLGLELIICGLEKQKCESGS 207

RESULT 4  
 W08297  
 ID W08297 standard; Protein; 207 AA.  
 AC W08297;  
 DT 19-MAR-1997 (first entry)



DE Mutated Tet repressor.  
KW Tet; tetracycline; Tc; operator; transcription; regulation; inducible;  
OS repressor; gene expression; therapy; transgenic animal; disease model.  
FH Synthetic.  
Key Location/Qualifiers  
FT misc\_difference 71 /label= substitution  
FT /note= "Lys replaces wild-type Glu"  
FT misc\_difference 95 /label= substitution  
FT /note= "Asn replaces wild-type Asp"  
FT misc\_difference 101 /label= substitution  
FT /note= "Ser replaces wild-type Leu"  
FT misc\_difference 102 /label= substitution  
FT /note= "Asp replaces wild-type Gly"  
PN W09640892-A1.  
PD 19-DEC-1996.  
PR 06-JUN-1996; U09049.  
PR 07-JUN-1995; US-485971.  
PA (BADI ) BASF AG.  
PI Bujard H, Gossen M, Helbl V, Hillen W, Schnappinger D;  
DR WPI; 97-052305/05.  
PT Nucleic acid encoding tetracycline-inducible transcription  
PT regulatory fusion protein - comprising modified tetracycline  
PT repressor able to bind mutant tet operator, fused to transcription  
PT regulator, useful for modulating eukaryotic gene expression  
PS Example 5; Page 86-87; 117pp; English.  
CC W08297 represents a mutated Tet repressor (TetR). This sequence was  
CC used to create a fusion protein, comprising the mutated TetR fused to a  
CC v-erbA silencer domain (see T45728, W08299). The fusion protein forms  
CC a Tc-controlled transcriptional activator (TRA). The main invention of  
CC the specification concerns modified TetR proteins that bind to modified  
CC class B tet operator sequences tetO-4C and tetO-6C (see T45711 and  
CC T48478). Modified tetR proteins can be fused to any transcription  
CC regulatory polypeptide and used to control transcription of a tetO-4C  
CC or tetO-6C linked gene. Nucleic acid encoding such a fusion protein  
CC may be introduced into a cell and transcription of the protein can be  
CC controlled by altering the concn. of tetracycline (or an analogue) in  
CC the cell, as appropriate. This ability to modulate gene expression in  
CC a predictable way is very useful in gene therapy and for recombinant  
CC protein prodn. in cultured cells or transgenic animals. The  
CC Tc-inducible system is also useful for the prodn. of transgenic animal  
CC models for the study of disease and also for the study of gene  
CC function e.g. during differentiation. The Tc-inducible system allows  
CC rapid activation of gene transcription without cellular toxicity, high  
CC concns. of inducer are not required.  
SQ Sequence 207 AA;

Query Match 100.0%; Score 1080; DB 1; Length 207;  
Best Local Similarity 100.0%; Pred. No. 3.4e-115;  
Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MSRLDKSVINSALELLNEVGIEGLTRKLAQKLGVEQPTLYWHVKNKRALDLAIAEML 60  
DB 1 MSRLDKSVINSALELLNEVGIEGLTRKLAQKLGVEQPTLYWHVKNKRALDLAIAEML 60  
QY 61 DRHHTHFCPLKGSQWDFLRNKAQKSFRCALLSHRNKAKVHSDTRPTEKQYETLENQALFL 120  
DB 61 DRHHTHFCPLKGSQWDFLRNKAQKSFRCALLSHRNKAKVHSDTRPTEKQYETLENQALFL 120  
QY 121 CQOGFSLENALYALSVAHGFTLGCVEQHQVAKEREETPTTDSMPPLLROAIELFDHQ 180  
DB 121 CQOGFSLENALYALSVAHGFTLGCVEQHQVAKEREETPTTDSMPPLLROAIELFDHQ 180  
QY 181 GAEPAPFLGLELLICGLEKOLKCESGS 207  
DB 181 GAEPAPFLGLELLICGLEKOLKCESGS 207  
RESULT 5  
W71327  
ID W71327 standard; Protein; 207 AA.  
AC W71327;  
DT 04-DEC-1998 (first entry)  
DE Protein encoded by mutant Tn10-derived tet repressor.

W71323  
ID W71323 standard; Protein; 335 AA.  
AC W71323;  
DT 04-DEC-1998 (first entry)  
DE Protein encoded by "reverse" Tet repressor sequence.  
KW Tet repressor; rTetR; tetracycline; regulation; expression;  
KW Tet operator-linked gene.  
OS Synthetic.  
PN US5814618-A.  
PD 29-SEP-1998.  
PR 07-JUN-1995; 485978.  
PR 07-JUN-1995; US-485978.  
PR 14-JUN-1993; US-076327.  
PR 14-JUN-1993; US-076726.  
PR 14-JUN-1994; US-260452.  
PR 01-JUL-1994; US-270637.  
PR 15-JUL-1994; US-275876.  
PR 06-FEB-1995; US-383754.  
PA (BADI ) BASF AG.  
PA (KNOL ) KNOLL AG.  
PI Bujard H, Gossen M;  
DR WPI; 98-541795/46.  
PT N-PSDB; V60076.  
PT Tetracycline based regulation of gene expression - uses a  
PT tetracycline operator sequence joined to a gene of interest, the  
PT gene of interest being induced in the presence, but not absence of  
PT the antibiotic  
PS Example 1; Columns 67-68; 63pp; English.  
CC The present sequence is encoded by a "reverse" Tet repressor (rTetR),  
CC which binds to its target DNA in the presence rather than the absence  
CC of tetracycline. The sequence was generated by chemical mutagenesis.  
CC rTetR is used in the course of the invention. The specification  
CC describes a method for regulating expression of a Tet (tetracycline)  
CC operator-linked gene in a cell of a subject. The method comprises  
CC introducing into the cell a nucleic acid encoding a fusion protein  
CC which inhibits transcription in eukaryotic cells, the fusion protein  
CC comprising a polypeptide which binds to a Tet operator sequence,  
CC operatively linked to heterologous second polypeptide which inhibits  
CC transcription in eukaryotic cells and modulating the concentration of  
CC tetracycline (analogue) in the subject. The method is used for the  
CC regulation of gene expression system, using tetracycline (analogues).  
CC The system enables a gene coupled to the system to be induced in the  
CC presence of Tet and then stopped when Tet is removed.  
SQ Sequence 335 AA;

Query Match 100.0%; Score 1080; DB 1; Length 335;  
Best Local Similarity 100.0%; Pred. No. 6.8e-115;  
Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MSRLDKSVINSALELLNEVGIEGLTRKLAQKLGVEQPTLYWHVKNKRALDLAIAEML 60  
DB 1 MSRLDKSVINSALELLNEVGIEGLTRKLAQKLGVEQPTLYWHVKNKRALDLAIAEML 60  
QY 61 DRHHTHFCPLKGSQWDFLRNKAQKSFRCALLSHRNKAKVHSDTRPTEKQYETLENQALFL 120  
DB 61 DRHHTHFCPLKGSQWDFLRNKAQKSFRCALLSHRNKAKVHSDTRPTEKQYETLENQALFL 120  
QY 121 CQOGFSLENALYALSVAHGFTLGCVEQHQVAKEREETPTTDSMPPLLROAIELFDHQ 180  
DB 121 CQOGFSLENALYALSVAHGFTLGCVEQHQVAKEREETPTTDSMPPLLROAIELFDHQ 180  
QY 181 GAEPAPFLGLELLICGLEKOLKCESGS 207  
DB 181 GAEPAPFLGLELLICGLEKOLKCESGS 207  
RESULT 6  
W71327  
ID W71327 standard; Protein; 207 AA.  
AC W71327;  
DT 04-DEC-1998 (first entry)  
DE Protein encoded by mutant Tn10-derived tet repressor.

KW Tet repressor; tetracycline; regulation; expression;  
 KW Tet operator-linked gene; tet operator.  
 OS Synthetic.  
 PN US5814618-A.  
 PD 29-SEP-1998.

PF 07-JUN-1995; 485978.  
 PR 07-JUN-1995; US-485978.  
 PR 14-JUN-1993; US-076327.  
 PR 14-JUN-1993; US-076726.  
 PR 14-JUN-1994; US-260452.  
 PR 01-JUL-1994; US-270637.  
 PR 15-JUL-1994; US-275876.  
 PR 06-FEB-1995; US-383754.  
 PA (BADI ) BASF AG.  
 PA (KNOL ) KNOLL AG.  
 PI Bujard H, Gossen M;  
 DR WPI: 98-541795/46.  
 DR N-PSDB; V60089.

PT Tetracycline based regulation of gene expression - uses a  
 PT tetracycline operator sequence joined to a gene of interest, the  
 PT gene of interest being induced in the presence, but not absence of  
 PT the antibiotic.  
 PS Claim 9; Columns 81-82; 63pp; English.  
 CC The present sequence is encoded by a mutant Tn10-derived tet repressor.  
 CC It is used in the course of the invention. The specification describes a  
 CC method for regulating expression of a Tet (tetracycline) operator-linked  
 CC gene in a cell of a subject. The method comprises introducing into the  
 CC cell a nucleic acid encoding a fusion protein which inhibits  
 CC transcription in eukaryotic cells, the fusion protein comprising a  
 CC polypeptide which binds to a tet operator sequence, operatively linked  
 CC to heterologous second polypeptide which inhibits transcription in  
 CC eukaryotic cells and modulating the concentration of a tetracycline  
 CC (analogue) in the subject. The method is used for the regulation of  
 CC gene expression system, using tetracycline (analogues). The system  
 CC enables a gene coupled to the system to be induced in the presence of  
 CC Tet and then stopped when Tet is removed.  
 SQ Sequence 207 AA;

Query Match 100.0%; Score 1080; DB 1; Length 207;  
 Best Local Similarity 100.0%; Pred. No. 3 4e-115;  
 Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSRLDKSVINSALLELNEVGIEGLTTRKLAQKLGVEOPTLYWHVKNKRALDLAIAEML 60  
 DB 1 MSRLDKSVINSALLELNEVGIEGLTTRKLAQKLGVEOPTLYWHVKNKRALDLAIAEML 60  
 QY 61 DRHHTFCPLKGESWQDFLRNKAQKSFRCALLSHRNGAKVHSDTRPTEKYETLENQALFL 120  
 DB 61 DRHHTFCPLKGESWQDFLRNKAQKSFRCALLSHRNGAKVHSDTRPTEKYETLENQALFL 120  
 QY 121 CQOGFSLENALYALSVAHGHTLGCVLDEQEHQVAKEREETPTTDSMPPLLRQAIEFLDHQ 180  
 DB 121 CQOGFSLENALYALSVAHGHTLGCVLDEQEHQVAKEREETPTTDSMPPLLRQAIEFLDHQ 180  
 QY 181 GAEPAPFLGLELIICGLEKQKCESGS 207  
 DB 181 GAEPAPFLGLELIICGLEKQKCESGS 207

RESULT 7

ID R64808 standard; Protein; 335 AA.  
 AC R64808; 1995 (first entry)  
 DT 17-JUL-1995  
 KW tTA transactivator.  
 KW tTA; transactivator; tetracycline-controllable transactivator;  
 KW conditional inactivation; homologous recombination; gene expression;  
 KW gene regulation; gene therapy; tetracycline-resistance; tetR;  
 KW transgenic animal.  
 OS Herpes simplex virus K12, KOS.  
 PN W09429442-A.  
 PD 22-DEC-1994.

PF 14-JUN-1994; U06734.  
 PR 14-JUN-1993; US-076327.  
 PA (BADI ) BASF AG.  
 PI Bujard H, Gossen M, Salfeld JG, Voss JW;  
 DR WPI: 95-036472/05.  
 DR N-PSDB; Q76264.

PT Regulatory systems using tetracycline-controllable transactivator  
 PT (tTA) - useful for conditional inactivation or modulation of  
 PT gene expression in a host cell or animal  
 PS Disclosure; Page 46-48; 103pp; English.  
 CC A 397 bp MluI/PokI fragment of pMSVP16 coding for the C-terminal 130  
 CC amino acids of VP16 of HSV was blunted and inserted in pUHD14-1.  
 CC The resulting plasmid, pUHD15-1, encoded a tetR-VP16 fusion protein,  
 CC or tetracycline-controllable transactivator (tTA), whose sequence is  
 CC given in Q76264, and the encoded protein sequence in R64808.  
 SQ Sequence 335 AA;

Query Match 98.0%; Score 1058; DB 1; Length 335;  
 Best Local Similarity 98.1%; Pred. No. 2.2e-112;  
 Matches 203; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MSRLDKSVINSALLELNEVGIEGLTTRKLAQKLGVEOPTLYWHVKNKRALDLAIAEML 60  
 DB 1 MSRLDKSVINSALLELNEVGIEGLTTRKLAQKLGVEOPTLYWHVKNKRALDLAIAEML 60  
 QY 61 DRHHTFCPLKGESWQDFLRNKAQKSFRCALLSHRNGAKVHSDTRPTEKYETLENQALFL 120  
 DB 61 DRHHTFCPLKGESWQDFLRNKAQKSFRCALLSHRNGAKVHSDTRPTEKYETLENQALFL 120  
 QY 121 CQOGFSLENALYALSVAHGHTLGCVLDEQEHQVAKEREETPTTDSMPPLLRQAIEFLDHQ 180  
 DB 121 CQOGFSLENALYALSVAHGHTLGCVLDEQEHQVAKEREETPTTDSMPPLLRQAIEFLDHQ 180  
 QY 181 GAEPAPFLGLELIICGLEKQKCESGS 207  
 DB 181 GAEPAPFLGLELIICGLEKQKCESGS 207

RESULT 8

ID R89763 standard; Protein; 207 AA.  
 AC R89763;  
 DT 07-JUL-1996 (first entry)  
 DE Wild type Tn-10 derived tet repressor.  
 KW Tet repressor; Herpes simplex virus; HSV; virion protein 16; VP16;  
 KW fusion protein; gene expression; regulation; inhibition; activation;  
 KW transcription.  
 OS Transposon Tn-10.  
 PN W09601313-A1.  
 PD 18-JAN-1996.  
 PF 29-JUN-1995; U08179.  
 PR 01-JUL-1994; US-270637.  
 PR 15-JUL-1994; US-275876.  
 PR 03-FEB-1995; US-383754.  
 PR 07-JUN-1995; US-486814.  
 PA (BUTA) BUJARD H.  
 PA (GOSSE) GOSSEN M.  
 PI Bujard H, Gossen M;  
 DR WPI: 96-087666/09.  
 DR N-PSDB; T11351.  
 PT New tetracycline-regulated transcription modulators - comprising  
 PT fusion proteins which bind to tet operator sequences to activate or  
 PT inhibit transcription  
 PS Claim 38; Page 78; 112pp; English.  
 CC Fusion proteins comprising a first polypeptide which binds to a tet  
 CC operator sequence in the presence of tetracycline or a tetracycline  
 CC analogue, operatively linked to a second polypeptide which either  
 CC activates or inhibits transcription in eukaryotic cells may be used  
 CC to activate or inhibit transcription. Such proteins may be used to  
 CC regulate gene expression in cells and may be particularly useful for  
 CC gene therapy and for expression of gene products in transgenic  
 CC organisms. Induction of gene expression is rapid, efficient and

CC strong, typically 1000-2000 fold. The inducing agent does not cause  
 CC pleiotropic effects or cytotoxicity in eukaryotic cells. This  
 CC sequence is the wild type in-10 tet repressor.  
 SQ Sequence 207 AA;

Query Match 98.0%; Score 1058; DB 1: Length 207;  
 Best Local Similarity 98.1%; Pred. No. 1.1e-112;  
 Matches 203; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MSRLDKSKVINSALLENVIGIEGLTRKLAQKLGVEQPTLYWHVKNKRALDLAIAEML 60  
 DB 1 MSRLDKSKVINSALLENVIGIEGLTRKLAQKLGVEQPTLYWHVKNKRALDLAIAEML 60

QY 61 DRHHTFCPLKGSWQDFLRNKAQKFRCALLSHRNKAKVHSDTRPTEKQYETLENQALFL 120  
 DB 61 DRHHTFCPLKGSWQDFLRNKAQKFRCALLSHRNKAKVHSDTRPTEKQYETLENQALFL 120

QY 121 COQGFSLNALYALSVAHGFTLGCVLDEQHOVAKEREETPTTDSMPPLLRQALIEFDHQ 180  
 DB 121 COQGFSLNALYALSVAHGFTLGCVLDEQHOVAKEREETPTTDSMPPLLRQALIEFDHQ 180

QY 181 GAEPALFGLLEIICGLERKQKCESGS 207  
 DB 181 GAEPALFGLLEIICGLERKQKCESGS 207

RESULT 9  
 W08296  
 ID W08296 standard; Protein; 207 AA.  
 AC W08296;  
 DE 19-MAR-1997 (first entry)  
 DE Wild-type E.coli Tn10-derived Tet repressor.  
 KW Tet; tetracycline; Tc; operator; transcription; regulation; inducible;  
 KW repressor; gene expression; therapy; transgenic animal; disease model;  
 KW Tn10; transposon 10.  
 OS Escherichia coli.  
 PN W09640892-Al.  
 PD 19-DEC-1996.  
 PF 06-JUN-1996; W09049.  
 PR 07-JUN-1995; US-485971.  
 PA (BADI) BASF AG.  
 PI Bujard H, Gossen M, Helbl V, Hillen W, Schnappinger D;  
 PI WPI; 97-052305/05.  
 PR Nucleic acid encoding tetracycline-inducible transcription  
 PT regulatory fusion protein - comprising modified tetracycline  
 PT repressor able to bind mutant tet operator, fused to transcription  
 PT regulator, useful for modulating eukaryotic gene expression  
 PS Example 4; Page 84-85; 117pp; English.  
 CC W08296 represents wild-type Tet repressor (TetR) protein, derived from  
 CC transposon 10 (Tn10) of E. coli. The wild-type sequence is used for  
 CC the production of a mutated TetR (e.g. so as to bind its target in the  
 CC presence rather than the absence of tetracycline). Mutant TetR proteins  
 CC can be fused to a transcriptional activator e.g. vpl6 (herpes simplex  
 CC virus virion protein 16), to form a Tc-controlled transcriptional  
 CC activator (tTA). The main invention of the specification concerns  
 CC modified TetR proteins that bind to modified class B tet operator  
 CC sequences tetO-4C and tetO-6C (see T45711 and T48478). Modified TetR  
 CC proteins can be fused to any transcription regulatory polypeptide and  
 CC used to control transcription of a tetO-4C or tetO-6C linked gene.  
 CC Nucleic acid encoding such a fusion protein may be introduced into a  
 CC cell and transcription of the protein can be controlled by altering the  
 CC concn. of tetracycline (or an analogue) in the cell, as appropriate.  
 CC This ability to modulate gene expression in a predictable way is very  
 CC useful in gene therapy and for recombinant protein prodn. in cultured  
 CC cells or transgenic animals. The Tc-inducible system is also useful for  
 CC the prodn. of transgenic animal models for the study of disease and  
 CC also for the study of gene function e.g. during differentiation. The  
 CC Tc-inducible system allows rapid activation of gene transcription.  
 CC Without cellular toxicity, high concns. of inducer are not required.  
 SQ Sequence 207 AA;

Query Match 98.0%; Score 1058; DB 1: Length 207;  
 Best Local Similarity 98.1%; Pred. No. 1.1e-112;  
 Matches 203; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MSRLDKSKVINSALLENVIGIEGLTRKLAQKLGVEQPTLYWHVKNKRALDLAIAEML 60  
 DB 1 MSRLDKSKVINSALLENVIGIEGLTRKLAQKLGVEQPTLYWHVKNKRALDLAIAEML 60

QY 61 DRHHTFCPLKGSWQDFLRNKAQKFRCALLSHRNKAKVHSDTRPTEKQYETLENQALFL 120  
 DB 61 DRHHTFCPLKGSWQDFLRNKAQKFRCALLSHRNKAKVHSDTRPTEKQYETLENQALFL 120

QY 121 COQGFSLNALYALSVAHGFTLGCVLDEQHOVAKEREETPTTDSMPPLLRQALIEFDHQ 180  
 DB 121 COQGFSLNALYALSVAHGFTLGCVLDEQHOVAKEREETPTTDSMPPLLRQALIEFDHQ 180

QY 181 GAEPALFGLLEIICGLERKQKCESGS 207  
 DB 181 GAEPALFGLLEIICGLERKQKCESGS 207

RESULT 10  
 W48357  
 ID W48357 standard; Protein; 651 AA.  
 AC W48357;  
 DT 20-JUL-1998 (first entry)  
 DE Multi-chimeric transactivating factor rTAER fusion protein.  
 KW Multi-chimeric transactivating factor; rTAER; tetR;  
 KW tetracycline repressor; HSV; oestrogen receptor; promoter;  
 KW packaging cell line; retrovirus; retroviral particle; vector;  
 KW gene delivery; gene therapy.  
 OS Chimeric - Escherichia coli.  
 OS Chimeric - Herpes simplex virus.  
 OS Chimeric - Mammalia.  
 PN W09805759-Al.  
 PD 12-FEB-1998.  
 PF 06-AUG-1997; U13846.  
 PR 07-AUG-1996; US-694652.  
 PA (CITY) CITY OF HOPE.  
 PA (REGC) UNIV CALIFORNIA.  
 PI Chen S, Friedmann T, Miyahara A, Prussak CE, Yee J;  
 PI WPI; 98-145602/13.  
 DR N-PSDB: V17756.  
 DR New packaging cell lines for pseudotyped retroviral vectors -  
 PT comprises sequences encoding retroviral Gag and Pol polypeptide(s)  
 PT and envelope protein  
 PS Disclosure: Fig 4A-C; 84pp; English.  
 CC This fusion protein comprises a multi-chimeric transactivator.  
 CC C-terminus) the Escherichia coli tetR polypeptide, the  
 CC designated rTAER, that is composed of (from the N-terminus to the  
 CC transcriptional activation domain of herpes simplex virus VP16, and  
 CC the ligand binding domain of the oestrogen receptor. An inducible  
 CC expression system of the invention is composed of rTAER and a  
 CC minimal promoter (see V17755) derived from the immediate early gene  
 CC of cytomegalovirus linked to 7 tandem copies of the tet operator  
 CC (tetO) that is the binding site for tetR, which in turn can be  
 CC linked to a nucleotide sequence of interest. The invention relates  
 CC to packaging cell lines derived from HeLa, D17, MDCK, BHK or  
 CC preferably Cf2Th cells and recombinant retroviral particles  
 CC produced by them, particularly pseudotyped retroviral particles.  
 CC Retroviral particles are produced by inducibly expressing an  
 CC envelope protein by linking an envelope protein-encoding nucleotide  
 CC sequence to the inducible expression system. The products can be  
 CC used for the inducible expression in cells of polypeptides, e.g.  
 CC cytotoxic products or therapeutic agents. The activation of the  
 CC inducible expression system requires 2 independent signals, thus  
 CC reducing the incidence of undesired transcriptional activation.  
 SQ Sequence 651 AA;

Query Match 98.0%; Score 1058; DB 1: Length 651;  
 Best Local Similarity 98.1%; Pred. No. 5.6e-112;  
 Matches 203; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MSRLDKSVINSALLELNEVGIEGLTRKLAQKLGVEQPTLYWHVKNKRALLDALAIEML 60  
 DB 1 MSRLDKSVINSALLELNEVGIEGLTRKLAQKLGVEQPTLYWHVKNKRALLDALAIEML 60  
 QY 61 DRHHTFCPLKGESWQDFLRNKAQKLGVEQPTLYWHVKNKRALLDALAIEML 120  
 DB 61 DRHHTFCPLKGESWQDFLRNKAQKLGVEQPTLYWHVKNKRALLDALAIEML 120  
 QY 121 CQGFSLNALYALSANGHFTLGCVLDEQHOVAKERETPTTDSMPPLRQAIELFDHQ 180  
 DB 121 CQGFSLNALYALSANGHFTLGCVLDEQHOVAKERETPTTDSMPPLRQAIELFDHQ 180  
 QY 181 GAEPALFGLLELIICGLEKQKLCESGS 207  
 DB 181 GAEPALFGLLELIICGLEKQKLCESGS 207

## RESULT 11

W47583  
 ID W47583 standard; Protein; 651 AA.  
 AC W47583; 1998 (first entry)  
 DT 21-JUL-1998 (first entry)  
 DE Multi-chimeric transactivator tTAER sequence.  
 KW Inducible expression system; modulation; cytotoxin; therapeutic;  
 KW tTAER; multi-chimeric; transactivator; tetR; Vp16;  
 KW ligand-binding domain.  
 OS Escherichia coli.  
 OS Herpes simplex virus.  
 PN W09805754-A2.  
 PD 12-FEB-1998.  
 PF 29-JUL-1997; U13221.  
 PR 07-AUG-1996; US-693940.  
 PA (CITY ) CITY OF HOPE.  
 PA (REGC ) UNIV CALIFORNIA.  
 PI Chen ST, Friedmann T, Yee JK;  
 PI WPI; 98-145597/13.  
 DR N-PSDB; V18690.  
 PT New inducible expression systems - comprising multi-chimeric  
 PT transactivator, induces transcription from promoter in the absence  
 PT of first ligand and presence of second ligand  
 PS Example 1; Fig 4; 72pp; English.  
 CC The sequence is that of multi-chimeric transactivator tTAER.  
 CC This is a fusion protein comprising, from N-terminus to C-terminus,  
 CC the E.coli tetr polypeptide, the transcription activation domain  
 CC of HSV Vp16 and the ligand-binding domain of an oestrogen receptor  
 CC from pHE14 (Kumar et al. 1986). DNA encoding it can be used as part of  
 CC an inducible expression system which induces transcription from the  
 CC promoter in the absence of the first ligand and presence of the  
 CC second. The products can be used for the inducible expression in cells  
 CC of polypeptides such as cytotoxic products or therapeutic products.  
 CC The activation of the inducible expression system requires 2  
 CC independent signals, which reduces the incidence of undesired  
 CC transcriptional activation.  
 SQ Sequence 651 AA;

Query Match 98.0%; Score 1058; DB 1; Length 651;  
 Best Local Similarity 98.1%; Pred. No. 5 6e-112;  
 Matches 203; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MSRLDKSVINSALLELNEVGIEGLTRKLAQKLGVEQPTLYWHVKNKRALLDALAIEML 60  
 DB 1 MSRLDKSVINSALLELNEVGIEGLTRKLAQKLGVEQPTLYWHVKNKRALLDALAIEML 60  
 QY 61 DRHHTFCPLKGESWQDFLRNKAQKLGVEQPTLYWHVKNKRALLDALAIEML 120  
 DB 61 DRHHTFCPLKGESWQDFLRNKAQKLGVEQPTLYWHVKNKRALLDALAIEML 120  
 QY 121 CQGFSLNALYALSANGHFTLGCVLDEQHOVAKERETPTTDSMPPLRQAIELFDHQ 180  
 DB 121 CQGFSLNALYALSANGHFTLGCVLDEQHOVAKERETPTTDSMPPLRQAIELFDHQ 180

QY 181 GAEPALFGLLELIICGLEKQKLCESGS 207  
 DB 181 GAEPALFGLLELIICGLEKQKLCESGS 207

## RESULT 12

W71326  
 ID W71326 standard; Protein; 207 AA.  
 AC W71326;  
 DT 04-DEC-1998 (first entry)  
 DE Protein encoded by wild type Tn10-derived tet repressor.  
 KW Tet repressor; tetracycline; regulation; expression;  
 KW Tet operator-linked gene; tet operator.  
 OS Unidentified.  
 PN US814618-A.  
 PD 29-SEP-1998.  
 PF 07-JUN-1995; 485978.  
 PR 07-JUN-1995; US-485978.  
 PR 14-JUN-1993; US-076327.  
 PR 14-JUN-1993; US-076726.  
 PR 14-JUN-1994; US-260452.  
 PR 01-JUL-1994; US-270637.  
 PR 15-JUL-1994; US-275876.  
 PR 06-FEB-1995; US-383754.  
 PA (BADI ) BASF AG.  
 PA (KNOL ) KNOLL AG.  
 PI Bujard H, Gossen M;  
 PI WPI; 98-541795/46.  
 DR N-PSDB; V60088.  
 PT Tetracycline based regulation of gene expression - uses a  
 PT tetracycline operator sequence joined to a gene of interest, the  
 PT gene of interest being induced in the presence, but not absence of  
 PT the antibiotic  
 PS Claim 4; Columns 77-80; 63pp; English.  
 CC The present sequence is encoded by wild type Tn10-derived tet repressor.  
 CC It is used in the course of the invention. The specification describes a  
 CC method for regulating expression of a Tet (tetracycline) operator-linked  
 CC gene in a cell of a subject. The method comprises introducing into the  
 CC cell a nucleic acid encoding a fusion protein which inhibits  
 CC transcription in eukaryotic cells, the fusion protein comprising a  
 CC polypeptide which binds to a Tet operator sequence, operatively linked  
 CC to heterologous second polypeptide which inhibits transcription in  
 CC eukaryotic cells and modulating the concentration of a tetracycline  
 CC (analogue) in the subject. The method is used for the regulation of  
 CC gene expression system, using tetracycline (analogues). The system  
 CC enables a gene coupled to the system to be induced in the presence of  
 CC Tet and then stopped when Tet is removed.  
 SQ Sequence 207 AA;

Query Match 98.0%; Score 1058; DB 1; Length 207;  
 Best Local Similarity 98.1%; Pred. No. 1.1e-112;  
 Matches 203; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MSRLDKSVINSALLELNEVGIEGLTRKLAQKLGVEQPTLYWHVKNKRALLDALAIEML 60  
 DB 1 MSRLDKSVINSALLELNEVGIEGLTRKLAQKLGVEQPTLYWHVKNKRALLDALAIEML 60  
 QY 61 DRHHTFCPLKGESWQDFLRNKAQKLGVEQPTLYWHVKNKRALLDALAIEML 120  
 DB 61 DRHHTFCPLKGESWQDFLRNKAQKLGVEQPTLYWHVKNKRALLDALAIEML 120  
 QY 121 CQGFSLNALYALSANGHFTLGCVLDEQHOVAKERETPTTDSMPPLRQAIELFDHQ 180  
 DB 121 CQGFSLNALYALSANGHFTLGCVLDEQHOVAKERETPTTDSMPPLRQAIELFDHQ 180  
 QY 181 GAEPALFGLLELIICGLEKQKLCESGS 207  
 DB 181 GAEPALFGLLELIICGLEKQKLCESGS 207

## RESULT 13

R48630

ID R48630 standard; Protein; 207 AA.  
AC R48630;  
DT 19-AUG-1994 (first entry)  
DE Sequence of the tetracycline repressor tetr.  
KW Tetracycline repressor; TetR; Tn10; transposon 10.  
OS Escherichia coli.  
PN WO9404672-A.  
PD 03-MAR-1994.  
PF 26-AUG-1993; U08230.  
PR 26-AUG-1992; US-935763.  
PA (DNX-) DNX CORP.  
PI Byrne G;  
DR WPI; 94-083191/10.  
DR N-PSDB; Q56710.  
PT Tetracycline repressor-mediated regulation system - useful for  
PT controlling gene expression in transgenic animals  
PS Disclosure: Page 41-42; 76pp; English.  
CC The inventors claim a construct which comprises an animal promoter  
CC element having a tetracycline repressor (tetR) operator. The  
CC promoter element may be the pPCK promoter which is tissue specific  
CC being expressed selectively in the liver and becoming active  
CC shortly prior to birth. The tetR sequence in the  
CC construct is 3' to a TATA-box sequence and is inserted into the NheI  
CC site of the pPCK promoter element. The entire sequence of the tetR  
CC repressor is given in Q56710/R48630.  
SQ Sequence 207 AA;

Query Match 97.5%; Score 1053; DB 1; Length 207;  
Best Local Similarity 97.6%; Pred. No. 4e-112;  
Matches 202; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
QY 1 MSRLDKSVINSALVELLNEVGIEGLTRKLAQKLGVEQPTLYWHVKNKRALDLAIEML 60  
DB 1 MSRLDKSVINSALVELLNEVGIEGLTRKLAQKLGVEQPTLYWHVKNKRALDLAIEML 60  
QY 61 DRHHTHFCPLKESWQDFLRNNAKSPRCALLSHRNGAKVHSDTRPTEKYETLENQALFL 120  
DB 61 DRHHTHFCPLKESWQDFLRNNAKSPRCALLSHRNGAKVHSDTRPTEKYETLENQALFL 120  
QY 121 COQGSLENALYALSVAHGHTLGCVLDEQHOVAKEREETPTTDSMPPLLRQAIEFLDHQ 180  
DB 121 COQGSLENALYALSVAHGHTLGCVLDEQHOVAKEREETPTTDSMPPLLRQAIEFLDHQ 180  
QY 181 GAEPAPFLGLELIIICGLEKQKCESGS 207  
DB 181 GAEPAPFLGLELIIICGLEKQKCESGS 207

RESULT 14  
R64809  
ID R64809 standard; Protein; 297 AA.  
AC R64809;  
DT 17-JUL-1995 (first entry)  
DE tTAS transactivator.  
KW tTA; tTAS; transactivator; tetracycline-controllable transactivator;  
KW conditional inactivation; homologous recombination; gene expression;  
KW gene regulation; gene therapy; tetracycline-resistance; tetR;  
KW transgenic animal.  
OS Herpes simplex virus K12, KOS.  
PN WO9429442-A.  
PD 22-DEC-1994.  
PF 14-JUN-1994; U06734.  
PR 14-JUN-1993; US-076327.  
PA (BADI) BASF AG.  
PI Bujard H, Gossen M, Salfeld JG, Voss JW;  
DR WPI: 95-036472/05.  
DR N-PSDB; Q76265.  
PT Regulatory systems using tetracycline-controllable transactivator  
PT (tTA) - useful for conditional inactivation or modulation of  
PT gene expression in a host cell or animal  
PS Disclosure: Page 50-51; 103pp; English.  
CC A DNA fragment of pMSVP16 coding for the C-terminal 97 amino

CC acids of VP16 of HSV was blunted and inserted in pUHD14-1. The  
CC resulting plasmid, pUHD151-1, encoded a tetr-VP16 fusion protein,  
CC or tetracycline-controllable transactivator (smaller version, tTAS),  
CC whose sequence is given in Q76265.  
SQ Sequence 297 AA;

Query Match 97.5%; Score 1053; DB 1; Length 297;  
Best Local Similarity 97.6%; Pred. No. 6.7e-112;  
Matches 202; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
QY 1 MSRLDKSVINSALVELLNEVGIEGLTRKLAQKLGVEQPTLYWHVKNKRALDLAIEML 60  
DB 1 MSRLDKSVINSALVELLNEVGIEGLTRKLAQKLGVEQPTLYWHVKNKRALDLAIEML 60  
QY 61 DRHHTHFCPLKESWQDFLRNNAKSPRCALLSHRNGAKVHSDTRPTEKYETLENQALFL 120  
DB 61 DRHHTHFCPLKESWQDFLRNNAKSPRCALLSHRNGAKVHSDTRPTEKYETLENQALFL 120  
QY 121 COQGSLENALYALSVAHGHTLGCVLDEQHOVAKEREETPTTDSMPPLLRQAIEFLDHQ 180  
DB 121 COQGSLENALYALSVAHGHTLGCVLDEQHOVAKEREETPTTDSMPPLLRQAIEFLDHQ 180  
QY 181 GAEPAPFLGLELIIICGLEKQKCESGS 207  
DB 181 GAEPAPFLGLELIIICGLEKQKCESGS 207

RESULT 15  
R85324  
ID R85324 standard; Protein; 297 AA.  
AC R85324;  
DT 13-MAR-1996 (first entry)  
DE tTAS transactivator.  
KW tTAS; transactivator; tetracycline-responsive promoter;  
KW gene expression; tetracycline-responsive promoter;  
KW transgenic animal.  
OS Chimeric Escherichia coli;  
OS Chimeric Herpes simplex virus.  
PN US5464738-A.  
PD 07-NOV-1995.  
PF 14-JUN-1993; 076726.  
PR 14-JUN-1993; US-076726.  
PA (BUJA/) BUJARD H.  
PA (GOSSE/) GOSSEN M.  
PI Bujard H, Gossen M;  
DR WPI: 95-392612/50.  
DR N-PSDB; T06868.  
PT Polynucleotide encoding trans:activator fusion protein contg. tet  
PT repressor - used to control expression of gene regulated by minimal  
PT promoter linked to tet operon, and vectors and cells where gene  
PT expression is regulated by tetracycline  
PS Disclosure: Fig 5A-B; 37pp; English.  
CC A fusion protein (R85324) composed of the Escherichia coli Tn10  
CC tet repressor (tetR) and a 127-amino acid C-terminal portion of the  
CC transcription activating domain of herpes simplex virus virion  
CC protein 16 (VP16) is the product of novel tetracycline-controlled  
CC transactivator tTA (T06868). The transactivator is used to regulate  
CC expression of a heterologous gene operably linked to a minimal promoter  
CC derived from human cytomegalovirus promoter IE (see T06869-70) and at  
CC least 1 tet operator (tetO) sequence. On/off regulation of expression  
CC of the heterologous protein by host eukaryotic cells is provided by  
CC varying the medium tetracycline conc. Transgenic animals producing  
CC a desired protein in their milk can also be produced.  
SQ Sequence 297 AA;

Query Match 97.5%; Score 1053; DB 1; Length 297;  
Best Local Similarity 97.6%; Pred. No. 6.7e-112;  
Matches 202; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
QY 1 MSRLDKSVINSALVELLNEVGIEGLTRKLAQKLGVEQPTLYWHVKNKRALDLAIEML 60  
DB 1 MSRLDKSVINSALVELLNEVGIEGLTRKLAQKLGVEQPTLYWHVKNKRALDLAIEML 60

```

Db 1 MSRLDKSVINSALLELNEVGIEGLTRKLAQKLGVEQPTLYWHVKNKRALLDALAIEML 60
QY 61 DRHHTHECPLKGESWQDFLRNKAQKSPFCALLSHRNGAKVHSDTRPTEKQYETILENOLAF 120
Db 61 DRHHTHECPLKGESWQDFLRNKAQKSPFCALLSHRNGAKVHSDTRPTEKQYETILENOLAF 120
QY 121 CQOGFSLLENALYALSAYGHFTLGCVLDEQEHQVAKEREETPTTDSMPPLLRQAIELFDHQ 180
Db 121 CQOGFSLLENALYALSAYGHFTLGCVLDEQEHQVAKEREETPTTDSMPPLLRQAIELFDHQ 180
QY 181 GAEPAPFLGLELIICGLEKOLKCESGS 207
Db 181 GAEPAPFLGLELIICGLEKOLKCESGS 207

```

Search completed: June 9, 1999, 12:59:17  
 Job time: 9196 sec

GenCore version 4.5  
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 9, 1999, 12:59:50 ; Search time 19.42 seconds  
(without alignments)  
399.294 Million cell updates/sec

Title: US-08-486-814-19  
Perfect score: 1080  
Sequence: 1 MSRLDKSVINSALELLNEV.....FGLELIICGLEKOLKCESGS 207

Scoring table: BLOSUM62

Searched: 116738 seqs, 37460341 residues

Database : PIR58:\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	DB	ID	Description
1	1053	97.5	207	1	RPECTN	repressor tetR - E
2	665	61.6	218	2	S07359	regulatory protein
3	665	61.6	218	2	S30287	regulatory protein
4	597	55.3	126	2	A26948	tetracycline resis
5	512	47.4	219	1	RPECYS	tetracycline repre
6	505.5	46.8	216	2	J01478	regulatory protein
7	505.5	46.8	216	2	S38655	tetR protein - Pse
8	493.5	45.9	216	1	RPECRL	tet repressor prot
9	493.5	45.9	211	2	I77569	tet repressor - Es
10	168.5	15.6	261	2	H70740	hypothetical prote
11	147.5	13.7	259	2	A40046	TetR repressor hom
12	123	11.4	191	2	A69900	hypothetical prote
13	106.5	9.9	210	2	S42419	probable transcrip
14	105.5	9.8	210	2	S42417	probable transcrip
15	103.5	9.6	192	2	C70487	transcription regu
16	102	9.4	189	2	H70042	transcription regu
17	99	9.2	196	2	B70827	hypothetical prote
18	96.5	8.9	246	2	E70861	hypothetical prote
19	94	8.7	179	2	B70391	transcription regu
20	93.5	8.7	236	2	S75298	hypothetical prote
21	92.5	8.6	236	2	D64855	hypothetical prote
22	91.5	8.5	196	2	D64918	glucuronide repres
23	91.5	8.5	194	2	C70035	transcription regu
24	91.5	8.5	213	2	F70946	probable regulato
25	86	8.0	205	2	S29308	hypothetical prote
26	84.5	7.8	543	2	S58095	hypothetical prote
27	83.5	7.7	963	1	A41919	kinesin heavy chai
28	83.5	7.7	216	2	C70649	hypothetical prote
29	83	7.7	195	2	S10899	regulatory protein
30	83	7.7	198	2	E69779	transcription regu
31	83	7.7	2748	2	S57976	nuclear migration
32	82.5	7.6	299	2	S39744	transcription regu
33	82.5	7.6	132	2	E69280	iron-dependent rep
34	82	7.6	212	2	C65165	tkk protein - Esch
35	82	7.6	365	2	S17885	Tcd37 protein - fr
36	81.5	7.5	2472	2	A35715	fodrin alpha chain
37	81	7.5	207	2	H70001	hypothetical prote
38	80.5	7.5	549	1	S19095	transcription fact
39	80.5	7.5	235	2	A64895	conserved hypothet

RESULT 1

RPECTN repressor tetR - Escherichia coli transposon Tn10

C:Species: Escherichia coli

C:Date: 17-Mar-1987 #sequence\_revision 17-Mar-1987 #text\_change 12-Jun-1998

C:Accession: A03576; S02667

R:Postle, K.; Nguyen, T.T.; Berstrand, K.P.

Nucleic Acids Res 12, 4849-4863, 1984

A:Title: Nucleotide sequence of the repressor gene of the Tn10 tetracycline resistance

A:Reference number: A03576; MUID:84247342

A:Accession: A03576

A:Molecule type: DNA

A:Residues: 1-207 <POS>

A:Cross-references: GB:X00694; NID:g43051; PID:g43052

R:Altschmid, L.; Baumeister, R.; Pfeleiderer, K.; Hillen, W.

EMBO J. 7, 4011-4017, 1988

A:Title: A threonine to alanine exchange at position 40 of tet repressor alters the r

A:Reference number: S02667; MUID:89091153

A:Accession: S02667

A:Status: not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-100 <ALT>

A:Cross-references: GB:J01830; NID:gl54845

C:Comment: This protein contains sequences similar to the DNA recognition regions of

C:Genetics:

A:Gene: tetR

C:Superfamily: tetracycline repressor

C:Keywords: antibiotic resistance; DNA binding; transcription regulation

F:25-46/Region: helix-turn-helix #status predicted

Query Match 97.5%; Score 1053; DB 1; Length 207;

Best Local Similarity 97.6%; Pred. No. 2.2e-91;

Matches 202; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MSRLDKSVINSALELLNEVIGLEGLTRKLAQKLGVEQPTLYWHVKNKRALDLAIEML 60

Db 1 MSRLDKSVINSALELLNEVIGLEGLTRKLAQKLGVEQPTLYWHVKNKRALDLAIEML 60

Qy 61 DRHHTFCPLKGSWQDFLRNKAHSFRCALLSHRNGAKVHSDTRTEKQYETLENQIAFL 120

Db 61 DRHHTFCPLKGSWQDFLRNKAHSFRCALLSHRNGAKVHSDTRTEKQYETLENQIAFL 120

Qy 121 CQGFSELENALYALSVAHGFTLGCVLDEQEHQVAKEREPTTDSMPPLLRQAIELEFDHQ 180

Db 121 CQGFSELENALYALSVAHGFTLGCVLDEQEHQVAKEREPTTDSMPPLLRQAIELEFDHQ 180

Qy 181 GAEPALFLGLELIICGLEKOLKCESGS 207

Db 181 GAEPALFLGLELIICGLEKOLKCESGS 207

RESULT 2

S07359

regulatory protein tetR - Escherichia coli plasmid RAI

N:Alternate names: tet repressor class D

C:Species: Escherichia coli

C:Date: 07-Sep-1990 #sequence\_revision 07-Sep-1990 #text\_change 12-Jun-1998

C:Accession: S07359

R:Unger, B.; Klock, G.; Hillen, W.

Nucleic Acids Res. 12, 7693-7703, 1984

A:Title: Nucleotide sequence of the repressor gene of the RAI tetracycline resistance





QY 61 DRHTHFCPLKGSWODFLRNKAKSFRCALLSHRNGAKVHSDTRPTKQYETLENQIAFL 120  
Db 61 TINTHTTPRDDDDWRSFLRNGACSFRRALLAYRDGARIHAGTRPAAPQMEKADAQURFL 120  
QY 121 COQGFSLLENALYALSAGVHFTLGCVLDEQEHQVAKERE---TPTTDSMP--PLLROAIEFLD 178  
Db 121 CDAGFSAGDATYALMALSYFTVGAVLEEQASDAERGEDQLTTSASTPARLOSAMKI 180  
QY 177 FDHGAEPAPFLFGLIICGLEK 199  
Db 181 VYEGGPDAAFERGLIIGGLEK 203

## RESULT 6

QJ1478  
regulatory protein tetr - Escherichia coli transposon Tn1721  
N;Alternate names: tetracycline resistance repressor  
C;Species: Escherichia coli  
C;Date: 17-Aug-1992 #sequence\_revision 17-Aug-1992 #text\_change 12-Jun-1998  
C;Accession: JQ1478; S02668; S24113  
R;Almeida, H.; Cresnar, B.; Greck, M.; Schmitt, R.  
Gene 111, 11-20, 1992  
A;Title: Complete nucleotide sequence of Tn1721: gene organization and a novel gene product  
A;Reference number: JQ1475; MUID:92192465  
A;Accession: JQ1478  
A;Molecule type: DNA  
A;Residues: 1-216 <ALL>  
A;Cross-references: EMBL:X61367; NID:g48194; PID:g48198  
A;Note: the authors translated the codon GAG for residue 19 as Asp and CAG for residue 8  
R;Altshied, L.; Baumeister, R.; Pfeleiderer, K.; Hillen, W.  
EMBO J. 7, 4011-4017, 1988  
A;Title: A threonine to alanine exchange at position 40 of tet repressor alters the repressor  
A;Reference number: S02667; MUID:89091153  
A;Accession: S02668  
A;Status: not compared with conceptual translation  
A;Molecule type: DNA  
A;Residues: 1-100 <ALT>  
C;Genetics:  
A;Gene: tetr  
C;Superfamily: tetracycline repressor  
F;25-46/Region: helix-turn-helix #status predicted

Query Match 46.8%; Score 505.5; DB 2; Length 216;  
Best Local Similarity 51.5%; Pred. No. 3.5e-40;  
Matches 104; Conservative 35; Mismatches 60; Indels 3; Gaps 2;

QY 1 MSRLDKSKVINSALLENEVGIEGLTTRKLAQKLGVEQPTLYWHVKNKRALLDALAEML 60  
Db 1 MTKLPQNTVIRAAADLLNEVGVDGLTTRKLAERLGVOOPALYWHFRNKRALLDALAEML 60  
QY 61 DRHTHFCPLKGSWODFLRNKAKSFRCALLSHRNGAKVHSDTRPTKQYETLENQIAFL 120  
Db 61 AENHTSVPRADDWRSFLGNARSFRQALLAYRDGARIHAGTRPGAPQMETADAQURFL 120  
QY 121 COQGFSLLENALYALSAGVHFTLGCVLDEQEHQVAKEREPTTDSMP--PLLROAIEFLD 178  
Db 121 CEAGFSAGDAVNALMTISYFTVGAVLEEQAGSDAGER-GGTVEQAPLSPLLRRAIDAFLD 179  
QY 179 HGAEPAPFLFGLIICGLEK 200  
Db 180 EAGPDAAFEQGLAVIVDGLAKR 201

## RESULT 7

S38655  
tet protein - Pseudomonas aeruginosa  
C;Species: Pseudomonas aeruginosa  
C;Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 12-Jun-1998  
C;Accession: S38655  
R;Trueman, P.; Sharpe, G.S.; Barth, P.T.  
submitted to the EMBL Data Library, November 1993

A;Reference number: S38655  
A;Accession: S38655  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-216 <TRU>  
A;Cross-references: EMBL:X75761; NID:g415984; PID:g415985  
C;Superfamily: tetracycline repressor  
F;25-46/Region: helix-turn-helix #status predicted

Query Match 46.8%; Score 505.5; DB 2; Length 216;  
Best Local Similarity 51.5%; Pred. No. 3.5e-40;  
Matches 104; Conservative 35; Mismatches 60; Indels 3; Gaps 2;

QY 1 MSRLDKSKVINSALLENEVGIEGLTTRKLAQKLGVEQPTLYWHVKNKRALLDALAEML 60  
Db 1 MTKLPQNTVIRAAADLLNEVGVDGLTTRKLAERLGVOOPALYWHFRNKRALLDALAEML 60  
QY 61 DRHTHFCPLKGSWODFLRNKAKSFRCALLSHRNGAKVHSDTRPTKQYETLENQIAFL 120  
Db 61 AENHTSVPRADDWRSFLGNARSFRQALLAYRDGARIHAGTRPGAPQMETADAQURFL 120  
QY 121 COQGFSLLENALYALSAGVHFTLGCVLDEQEHQVAKEREPTTDSMP--PLLROAIEFLD 178  
Db 121 CEAGFSAGDAVNALMTISYFTVGAVLEEQAGSDAGER-GGTVEQAPLSPLLRRAIDAFLD 179  
QY 179 HGAEPAPFLFGLIICGLEK 200  
Db 180 EAGPDAAFEQGLAVIVDGLAKR 201

## RESULT 8

RPECRI  
tet repressor protein (Tn 1721) - Escherichia coli plasmid RP1  
C;Species: Escherichia coli  
C;Date: 25-Feb-1985 #sequence\_revision 25-Feb-1985 #text\_change 12-Jun-1998  
C;Accession: A03574  
R;Waters, S.H.; Rogowsky, P.; Grinstad, J.; Altenbuchner, J.; Schmitt, R.  
Nucleic Acids Res. 11, 6089-6105, 1983  
A;Title: The tetracycline resistance determinants of RP1 and Tn1721: nucleotide sequence  
A;Reference number: A93486; MUID:83299270  
A;Accession: A03574  
A;Molecule type: DNA  
A;Residues: 1-216 <WAT>  
A;Cross-references: GB:X00006; NID:g42508; PID:g42509  
C;Genetics:  
A;Gene: tetr  
A;Genome: plasmid  
C;Superfamily: tetracycline repressor  
C;Keywords: antibiotic resistance; DNA binding; transcription regulation  
F;25-46/Region: helix-turn-helix #status predicted

Query Match 45.9%; Score 495.5; DB 1; Length 216;  
Best Local Similarity 50.5%; Pred. No. 3.1e-39;  
Matches 102; Conservative 36; Mismatches 61; Indels 3; Gaps 2;

QY 1 MSRLDKSKVINSALLENEVGIEGLTTRKLAQKLGVEQPTLYWHVKNKRALLDALAEML 60  
Db 1 MTKLPQNTVIRAAADLLNEVGVDGLTTRKLAERLGVOOPALYWHFRNKRALLDALAEML 60  
QY 61 DRHTHFCPLKGSWODFLRNKAKSFRCALLSHRNGAKVHSDTRPTKQYETLENQIAFL 120  
Db 61 AENHTSVPRADDWRSFLTNARSFRQALLAYRDGARIHAGTRPGAPQMETADAQURFL 120  
QY 121 COQGFSLLENALYALSAGVHFTLGCVLDEQEHQVAKEREPTTDSMP--PLLROAIEFLD 178  
Db 121 CEAGFSAGDAVNALMTISYFTVGAVLEEQAGSDSEGER-GGTVEQAPLSPLLRRAIDAFLD 179  
QY 179 HGAEPAPFLFGLIICGLEK 200  
Db 180 EAGPDAAFEQGLAVIVDGLAKR 201

```
RESULT 9
177569
tet repressor - Escherichia coli
C:Species: Escherichia coli
C:Date: 07-Jun-1996 #sequence_revision 07-Jun-1996 #text_change 12-Jun-1998
C:Accession: I77569
R:Tovar, K.; Ernst, A.; Hillen, W.
Mol. Gen. Genet. 215, 76-80, 1988
A:Title: Identification and nucleotide sequence of the class E tet regulatory elements
A:Reference number: I57766; MUID:89201249
A:Accession: I77569
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-211 <RES>
A:Cross-references: EMBL:X14035; NID:943055; PID:g43056
C:Superfamily: tetracycline repressor
I:45-46/Region: helix-turn-helix #status predicted

Query Match 45.7%; Score 493.5; DB 2; Length 211;
Best Local Similarity 50.2%; Pred. No. 4.6e-39;
Matches 101; Conservative 29; Mismatches 68; Indels 3; Gaps 1;

QY 1 MSRLDKVINSALLENEVGIEGLTRKLAQKLGVEOPTLYWHVKNKRALDLAIAEML 60
D: 1 MARLSLDDVISMAFLTLDSEGLTTRKLAQSLKIEOPTLYWHVKNKQTLMMNLSEAIL 60
QY 61 DRHHTHFCPLKGSQWDFLRNKAFCRALLSHRNKAKVHSDTRPTKEQYETLENQIAFL 120
D: 61 AKHHTSRAPLPTESWQOFLQENALSFRRKALLVHRDGLHGTSTPTPPQFEQAQLRCL 120
QY 121 COQGSLENALYALSAVGHFTLGCVLDEQEHQVAKEREPTTDSMPPLLRQAIELEFDHQ 180
D: 121 CDAGSVEALFLOSISGHFTLGALEE---QATNQIENNHIIVDAAPPLQEAENIQART 177
QY 181 GAEPALFGLLEIICGLEKQL 201
D: 178 SAEMAFHFLKSLIFGSAQL 198

RESULT 10
H70740
Hypothetical protein Rv1353c - Mycobacterium tuberculosis (strain H37Rv)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 17-Jul-1998
C:Accession: H70740
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Nature 393, 537-544, 1998
A:Authors: Krogh, A.; McLean, J.; Moule, S.; Murphy, L.; Oliver, S.; Osborne, J.; Quail, A.; Taylor, K.; Whitehead, S.; Barrett, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A:Reference number: A70500; MUID:98295987
A:Accession: H70740
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-261 <COL>
A:Cross-references: GB:275555; GB:AL123456; NID:93261608; PID:e250355; PID:gl419056
A:Experimental source: strain H37Rv
C:Genetics:
A:Gene: Rv1353c

Query Match 15.68; Score 168.5; DB 2; Length 261;
Best Local Similarity 27.08; Pred. No. 1.5e-08;
Matches 61; Conservative 32; Mismatches 78; Indels 55; Gaps 6;

QY 4 LDKSKVINSALLENEVGIEGLTRKLAQKLGVEOPTLYWHVKNKRALDLAIAEMLDRH 63
D: 16 INPEDIISGAELAAQVSDINLMPKLGKLVGVTSYIYFRKDKDLNATDRALSKY 75
QY 64 HTHFCPLKGSQWDFLRNKAFCR---CALL-----SHRNKAKVHSDTRP 105

RESULT 11
A40046
Tet repressor homolog actII-1 - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 12-Jun-1998
C:Accession: A40046
R:Fernandez-Moreno, M.A.; Caballero, J.L.; Hopwood, D.A.; Malpartida, F.
Cell 66, 769-780, 1991
A:Title: The act cluster contains regulatory and antibiotic export genes, direct target
A:Reference number: A40046; MUID:91347376
A:Accession: A40046
A:Molecule type: DNA
A:Residues: 1-259 <FER>
A:Cross-references: GB:M64683; NID:g153143; PID:g455360
C:Genetics:
A:Start codon: GTG
F:52-73/Region: helix-turn-helix #status predicted

Query Match 13.7%; Score 147.5; DB 2; Length 259;
Best Local Similarity 26.0%; Pred. No. 1.3e-06;
Matches 57; Conservative 35; Mismatches 90; Indels 37; Gaps 8;

QY 4 LDKSKVINSALLENEVGIEGLTRKLAQKLGVEOPTLYWHVKNKRALDLAIAEML 57
D: 31 LTQDRIVVTALGILDAEGLDALSRRRLAQELKTHGASLYAHVGNRDELDDLVDIVLTV 90
QY 58 EMLDRHHTHFCPLKGSQWDFLRNKAFCRALLSHRNKAKVHSDTRPT-KEQYETLENQ 116
D: 91 EYPE-----PEPGR-WAEQVKEMCRSLRMLRAHLDLARIADRVPLGPNMGVMGERT 142
QY 117 LAFLCQOQGSLENALYALSAVGHFTLGCVLDEQEHQVAKER-----RET 160
D: 143 MNLSSGGLHDELAAYGGDLISFTVTAELQSSRNPTQEGQAGVAFDOLHGKYLKSL 202
QY 161 PTTDSMPPLLRQA---IELFDHQAEPALFGLLEIICGL 197
D: 203 PAT-SFPNLVHLAAGPITSLD---SDRRFELGLEIILAGL 237

RESULT 12
A69900
Hypothetical protein yobS - Bacillus subtilis
C:Species: Bacillus subtilis
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 24-Sep-1998
C:Accession: A69900
R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Ber
C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.;
A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabbet, C.; Ferrari,
Nature 390, 249-256, 1997
A:Authors: Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galleron, N.; Gh
wood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.; Itaya, M
Konigstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois, S.; Laube
A:Authors: Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maugel, C.; Med
, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle, D.; Por
Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon, E.; Schl
A:Authors: Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Sroror, S.J.; Ser
anakhshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama, S.; Vanden
anahat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.; Yoshikaw
A:Authors: Zumstein, E.; Yoshikawa, H.; Danchin, A.
A:Title: The complete sequence of the Gram-positive bacterium Bacillus subtilis
```

S4241/  
 probable transcription repressor mtrR - Neisseria gonorrhoeae (strain CH95)  
 C.Species: Neisseria gonorrhoeae  
 A.Variety: strain CH95  
 C.Date: 07-May-1998 #sequence\_revision 15-May-1998 #text\_change 18-Sep-1998

Mol. Microbiol. 11, 769-775, 1994  
A:Title: Regulation of the permeability of the gonococcal cell envelope by the mtr sy  
A:Reference number: S42417  
A:Accession: S42417  
A:Status: nucleic acid sequence not shown  
A:Molecule type: DNA  
A:Residues: 1-210 <PAN>  
A:Cross-references: EMBL:Z25796; NID:g438130; PID:g438191  
A:Experimental source: CH95  
C:Genetics:  
A:Gene: mtrR  
C:Superfamily: probable transcription repressor mtrR  
C:Keywords: DNA binding; transcription regulation

Query Match 9.8%; Score 105.5; DB 2; Length 210;  
Best Local Similarity 26.8%; Pred. NO. 0.0086;  
Matches 51: Conservative 23; Mismatches 73; Indels 43; Gaps 8;

QY	6	KSKVINSALELLNEV	IEGLTTRFKLAKGKIQE	PTLYHTVKNKRALD	ALATEMLDRHHT	65	
Db	12	KEHMLAAUETFR	KGIARTSLSNEIAQAAG	VTFGALYWHFKKED	LDALFORICDDIEN	71	
QY	66	----	HFCPLKGESWQD	FLRNKAKSFRCALL	SHRNKAGVHSDTR	PTPEKQVETLENOLAFIC	121
Db	72	CIAQDAADABGG	SWTVF-RHTLLHFFER	QNSNDIYYKFH-	-----	NILFKC	116
QY	122	QOGFSLNAYALSA	VSGHFTTGLCVLEDO	QHOVAKERETPT	TDSDMPLPQRAIE	---LFD	178
Db	117	FR--TFQNA--	AVIATA-----	RKHOATWREKIT	-----	AVTEAVENODLAD	155

```

RESULT 15
C70487
transcription regulator TetR/AcrR family - Aquifex aeolicus
t:Species: Aquifex aeolicus
C:Status: 08-May-1998 #sequence_revision 08-May-1998 #text_change
C:Accession: C70487
Y:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.;

```

A:Accession: C/046/  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-192 <AQF>  
A:Cross-references: GB:AE000776; NID:g2984355; PID:g2984362; GB:AE000657  
A:Experimental source: strain VF5  
A:Genetics: C:Genetics:  
A:Gene: acrR2

Query Match 9.6%; Score 103.5; DB 2; Length 192;  
Best Local Similarity 22.7%; Pred. No. 0.012;  
Matches 45; Conservative 35; Mismatches 67; Indels 51; Gaps 8;  
Qy 6 KSKVINSALELNEVGISGLTRKLAQKLGVEQPTLYWHVRNK----RALLDALAIEMLD 61



P08238 homo sapien  
P75342 mycoplasma

44 73 5 6 8 723 1 H59B\_HUMAN  
45 73 6 8 1244 1 Y307\_MYCPN

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 9, 1999, 13:01:02 ; Search time 16.52 Seconds  
(without alignments)  
336.316 Million cell updates/sec

Title: US-08-486-814-19  
Perfect score: 1080  
Sequence: 1 MSRLDKSKVNSALELLNEV.....FGLELICGLEKOLKCESGS 207

Scoring table: BLOSUM62

Searched: 74019 seqs, 26840295 residues

Database : SwissProt\_36.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1053	97.5	207	1	TER2_ECOLI	P04483 escherichia
2	665	61.6	218	1	TER8_PASPI	P51562 pasteurella
3	663	61.4	207	1	TER8_PASPI	P51561 pasteurella
4	660	61.1	217	1	TER4_ECOLI	P09164 escherichia
5	512	47.4	219	1	TER3_ECOLI	P03039 escherichia
6	505.5	46.8	216	1	TER1_ECOLI	P03038 escherichia
7	493.5	45.7	211	1	TER5_ECOLI	P21337 escherichia
8	473	43.8	210	1	TER7_VIBAN	P51560 vibrio angu
9	168.5	15.6	261	1	Y07H_MYCTU	Q11023 mycobacteri
10	106.5	9.9	210	1	MYTH_NEIGO	P39897 neisseria g
11	92.5	8.6	236	1	YCFQ_ECOLI	P75952 escherichia
12	91.5	8.5	196	1	YDUR_ECOLI	Q59431 escherichia
13	86	8.0	955	1	PHSL_IPOBA	P27598 ipomea bat
14	84.5	7.8	577	1	THL1_SCHPO	Q09684 schizosacch
15	83.5	7.7	963	1	KINH_HUMAN	P33176 homo sapien
16	83	7.7	195	1	BETI_ECOLI	P17446 escherichia
17	82.5	7.6	299	1	YWFK_BACSU	P39647 bacillus su
18	82	7.6	212	1	TTK_ECOLI	P06969 escherichia
19	81.5	7.5	2472	1	SPCN_HUMAN	Q13813 homo sapien
20	80.5	7.5	549	1	CFIA_DROME	P16241 drosophila
21	80	7.4	188	1	YP23_STAAG	P23217 staphylococ
22	79.5	7.4	1391	1	N157_YEAST	P40064 saccharomyc
23	79	7.3	187	1	Y893_HAEIN	P44923 haemophilus
24	78.5	7.3	770	1	PRTP_VZVD	P09284 varicella-z
25	78.5	7.3	218	1	TTK_HAEIN	P29280 haemophilus
26	78	7.2	215	1	ACRR_ECOLI	P34000 escherichia
27	78	7.2	519	1	N1FL_AZOVI	P30663 azotobacter
28	77.5	7.2	212	1	YDCD_ECOLI	P75899 escherichia
29	77	7.1	192	1	BM3R_BACME	P43506 bacillus me
30	77	7.1	2748	1	NUM1_YEAST	Q00402 saccharomyc
31	77	7.1	673	1	PLZF_HUMAN	Q05516 homo sapien
32	76.5	7.1	838	1	PHSH_SOLTU	P32811 solanum tub
33	75	6.9	1324	1	CUT3_SCHPO	P41004 schizosacch
34	75	6.9	2319	1	FAB_MOUSE	Q06194 mus musculu
35	75	6.9	1509	1	MYSN_ACACA	P05659 acanthamoeb
36	74.5	6.9	557	1	ATRA_ECOLI	P03959 escherichia
37	74	6.9	2133	1	FAB_PIG	P12263 sus scrofa
38	74	6.9	474	1	LAM3_MOUSE	P48680 mus musculu
39	74	6.9	2116	1	MY52_DICDI	P08799 dictyosteli
40	74	6.9	892	1	YJ06_YEAST	P40985 saccharomyc
41	74	6.9	1207	1	YL08_CAEEL	P34402 caenorhabdi
42	73.5	6.8	323	1	FRA2_CHICK	P18625 gallus gall
43	73.5	6.8	728	1	HS9A_CHICK	P11501 gallus gall

## ALIGNMENTS

```

RESULT 1
TER2_ECOLI
ID TER2_ECOLI STANDARD; PRT; 207 AA.
AC P04483;
DT 13-AUG-1987 (REL. 05, CREATED)
DT 13-AUG-1987 (REL. 05, LAST SEQUENCE UPDATE)
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE TETRACYCLINE REPRESSOR PROTEIN CLASS B (TRANSPONSON TN10).
GN TETR.
OS ESCHERICHIA COLI.
OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;
OC ENTEROBACTERIACEAE.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 84247342.
RA POSTLE K., NGUYEN T.T., BERTRAND K.P.;
RL NUCLEIC ACIDS RES. 12:4849-4863(1984).
RN [2]
RP SEQUENCE OF 1-22 FROM N.A.
RX MEDLINE; 84005886.
RA BERTRAND K.P., POSTLE K., WRAY L.V. JR., REZNIKOFF W.S.;
RL GENE 23:149-156(1983).
RN [3]
RP MUTAGENESIS.
RX SMITH L.D., BERTRAND K.P.;(1988).
RL J. MOL. BIOL. 203:949-959(1988).
CC -1- FUNCTION: TETR IS THE REPRESSOR OF THE TETRACYCLINE RESISTANCE
CC ELEMENT; ITS AMINO-TERMINAL REGION FORMS A HELIX-TURN-HELIX
CC STRUCTURE AND BINDS DNA. BINDING OF TETRACYCLINE TO TETR REDUCES
CC THE REPRESSOR AFFINITY FOR THE TETRACYCLINE RESISTANCE GENE
CC (TETA) PROMOTER OPERATOR SITES.
CC -1- SIMILARITY: BELONGS TO THE TETR/ACRR FAMILY OF TRANSCRIPTIONAL
CC REGULATORS.
DR EMBL; X00694; G43052; -.
DR EMBL; J01830; G154846; -.
DR PIR; A03576; RPECTN.
DR PROSITE; PS01081; HTH_TETR_FAMILY; 1.
DR TRANSCRIPTION REGULATION; REPRESSOR; TRANSPOSABLE ELEMENT;
KW DNA-BINDING; ANTIBIOTIC RESISTANCE.
FT DNA_BIND 26 45 H-T-H MOTIF (POTENTIAL).
FT SITE 64 64 INVOLVED IN BINDING TO [MG-TC]+ (BY
FT METAL 100 100 MAGNESIUM (OF [MG-TC]+ COMPLEX) (BY
FT MUTAGEN 64 64 SIMILARITY).
FT MUTAGEN 82 82 H->Y: REDUCES AFFINITY FOR TETRACYCLINE
FT MUTAGEN 82 82 N->H: REDUCES AFFINITY FOR TETRACYCLINE
FT MUTAGEN 103 103 T->I: REDUCES AFFINITY FOR TETRACYCLINE
FT SEQUENCE 207 AA; 23354 MW; 4AB4DD5A CRC32;

```

Query Match 97.5%; Score 1053; DB 1; Length 207;  
Best Local Similarity 97.6%; Pred. No. 3 4e-92;  
Matches 202; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

```

Oy 1 MSRLDKSKVNSALELLNEVGIEGLTTRKLAQKLGVEQPTLYWHVKNKRALLDALATEML 60
Db 1 MSRLDKSKVNSALELLNEVGIEGLTTRKLAQKLGVEQPTLYWHVKNKRALLDALATEML 60
Oy 61 DRHHTFCPLKGSWQDFLRNNAKSFRCALLSHRNGAKVHSDTRPTKQYETLENQALFL 120
Db 61 DRHHTFCPLKGSWQDFLRNNAKSFRCALLSHRNGAKVHSDTRPTKQYETLENQALFL 120

```



NUCLEIC ACIDS RES. 12:7693-7703(1984).  
(2)  
SEQUENCE FROM N.A.  
SPECIES-S ORDONEZ; STRAIN-BM2000; PLASMID-PIP173;  
MEDLINE; 93204906.  
ALLARD J.D., GIBSON M.L., VU L.H., NGUYEN T.T., BERTRAND K.P.;  
MOL. GEN. GENET. 237:301-305(1993).  
(3)  
X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS), AND PARTIAL SEQUENCE.  
MEDLINE; 94204640.  
HINRICHS W., KISKER C., DUEVEL C., MUELLER A., TOVAR K., HILLEN W.,  
SAENGER W.;  
SCIENCE 264:418-420(1994).  
-!- FUNCTION: TETR IS THE REPRESSOR OF THE TETRACYCLINE RESISTANCE  
ELEMENT; ITS AMINO-TERMINAL REGION FORMS A HELIX-TURN-HELIX  
STRUCTURE AND BINDS DNA. BINDING OF TETRACYCLINE TO TETR REDUCES  
THE REPRESSOR AFFINITY FOR THE TETRACYCLINE RESISTANCE GENE  
(TETA) PROMOTER OPERATOR SITES.  
-!- SUBUNIT: HOMODIMER.  
-!- INDUCTION: BY THE [MG-TETRACYCLINE]+ COMPLEX.  
-!- SIMILARITY: BELONGS TO THE TETR/ACR FAMILY OF TRANSCRIPTIONAL  
REGULATORS.  
EMBL; X01083; G43054; -;  
EMBL; X5876; G49075; -;  
PIR; S07359; S07359.  
PIR; S30287; S30287.  
PDB; 2TCT; 03-APR-96.  
PDB; 2TRT; 20-JUN-96.  
PROSITE; PS01081; HTH\_TETR\_FAMILY; 1.  
TRANSCRIPTION REGULATION; REPRESSOR; DNA-BINDING; PLASMID;  
ANTIBIOTIC RESISTANCE; 3D-STRUCTURE.  
INIT\_MET 0  
DNA\_BIND 0  
SITE 25 44 H-T-H MOTIF (POTENTIAL).  
SITE 63 63 INVOLVED IN BINDING TO [MG-TC]+.  
METAL 99 99 MAGNESIUM (OF [MG-TC]+ COMPLEX).  
SEQUENCE 217 AA; 24287 MW; 830EF8EB CRC32;  
Query Match 61.1%; Score 660; DB 1; Length 217;  
Best Local Similarity 63.5%; Pred. No. 3.5e-55;  
Matches 127; Conservative 27; Mismatches 46; Indels 0; Gaps 0;  
OY 2 SRLDKSVINSALELLNEVEGIEGLTTRKLAQKLGVEOPTLYWHVKNKRALDLALAEML 61  
DB 1 ARLNRESVIDAALLNETGIDGTTTRKLAQKLGIEOPTLYWHVKNKRALDLALAEML 60  
OY 62 RHHTHFCPLKGSQWDFLRNKAQKSFRCALLSHRNGAKVHSDTRPTEKQYETLENQALFL 121  
DB 61 RHHDYSLPAAGESQWDFLRNKAQKSFRCALLSHRNGAKVHSDTRPTEKQYETLENQALFL 120  
OY 122 QOQFSLNALVALSAVGHFTLGCVLDEQHOVAKEERETPTTDSMPPLLRQALFLDHOG 181  
DB 121 ENGFSRLDGLYSAISAVSHFTLGVLEQOQHTAALTDRAAPDENLPLLRQALQIMDSDD 180  
OY 182 AEPALFGLLELIGCLEKOL 201  
DB 181 GEQAPLHGLSLINGFEVOL 200  
RESULT 5  
TER3\_ECOLI STANDARD; PRT; 219 AA.  
AC P03039;  
DT 21-JUL-1986 (REL. 01, CREATED)  
DT 21-JUL-1986 (REL. 01, LAST SEQUENCE UPDATE)  
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)  
DE TETRACYCLINE REPRESSOR PROTEIN CLASS C.  
GN TETR.  
OS ESCHERICHIA COLI.  
OC PLASMID PSC101.  
OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;  
OC ENTEROBACTERIACEAE.  
RN [1]

SEQUENCE FROM N.A.  
MEDLINE; 88216101.  
BROW M.A.D., PESIN R., SUTCLIFFE J.G.;  
MOL. BIOL. EVOL. 2:1-12(1985).  
-!- FUNCTION: TETR IS THE REPRESSOR OF THE TETRACYCLINE RESISTANCE  
ELEMENT; ITS AMINO-TERMINAL REGION FORMS A HELIX-TURN-HELIX  
STRUCTURE AND BINDS DNA. BINDING OF TETRACYCLINE TO TETR REDUCES  
THE REPRESSOR AFFINITY FOR THE TETRACYCLINE RESISTANCE GENE  
(TETA) PROMOTER OPERATOR SITES.  
-!- SIMILARITY: BELONGS TO THE TETR/ACR FAMILY OF TRANSCRIPTIONAL  
REGULATORS.  
EMBL; M36272; G150946; -;  
PIR; A03575; RPECYS.  
PROSITE; PS01081; HTH\_TETR\_FAMILY; 1.  
TRANSCRIPTION REGULATION; REPRESSOR; DNA-BINDING; PLASMID;  
ANTIBIOTIC RESISTANCE.  
DNA\_BIND 26 45 H-T-H MOTIF (POTENTIAL).  
SITE 64 64 INVOLVED IN BINDING TO [MG-TC]+ (BY  
SIMILARITY).  
METAL 100 100 MAGNESIUM (OF [MG-TC]+ COMPLEX) (BY  
SIMILARITY).  
SEQUENCE 219 AA; 24174 MW; 6C9E1828 CRC32;  
Query Match 47.4%; Score 512; DB 1; Length 219;  
Best Local Similarity 51.7%; Pred. No. 3e-41;  
Matches 105; Conservative 33; Mismatches 61; Indels 4; Gaps 1;  
OY 1 MSRLDKSVINSALELLNEVEGIEGLTTRKLAQKLGVEOPTLYWHVKNKRALDLALAEML 60  
DB 1 MNKLOREAVIRTALELLNDVGMGLTTRLAERLGVQPALYWHFKNKRALDLALAEML 60  
OY 61 DRHHTHFCPLKGSQWDFLRNKAQKSFRCALLSHRNGAKVHSDTRPTEKQYETLENQALFL 120  
DB 61 TINHTHPTDRDDWRSEFLKGNACSFRRALLAYRDGARIHAGTRPAAPQMEKADQURFL 120  
OY 121 COQFSLNALVALSAVGHFTLGCVLDEQHOVAKEERE---TPTTDSMPPLLRQALFL 176  
DB 121 CDAGFSAGDATTALMAISYFTVGAVLEQOQASEADAERGEDOLTTASTMPARLQSAWKI 180  
OY 177 FDHQAEPALFGLLELIGCLEK 199  
DB 181 VYEGGPDAAFERGLALIIGGLEK 203  
RESULT 6  
TER1\_ECOLI STANDARD; PRT; 216 AA.  
AC P03038;  
DT 21-JUL-1986 (REL. 01, CREATED)  
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)  
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)  
DE TETRACYCLINE REPRESSOR PROTEIN CLASS A (TRANSPONSON 1721).  
GN TETR.  
OS ESCHERICHIA COLI.  
OC PLASMID RPI AND PLASMID RP4.  
OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;  
OC ENTEROBACTERIACEAE.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TRANSPONSON-TN1721;  
RX MEDLINE; 92192465.  
RA ALLMEIER H., CRESNAR B., GRECK M., SCHMITT R.;  
RL GENE 111:11-20(1992).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC PLASMID-PLASMID RP4;  
RA TRUENMAN P., SHARPE G.S., BARTH P.T.;  
RL SUBMITTED (NOV-1993) TO EMBL/GENBANK/DBJ DATA BANKS.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC TRANSPONSON-TN1721; PLASMID-RP1;  
RX MEDLINE; 83299270.

RA WATERS S.H., ROGOWSKY P., GRINSTED J., ALTENBUCHNER J., SCHMITT R.;  
 RL NUCLEIC ACIDS RES. 11:6089-6105(1982).  
 CC -!- FUNCTION: TETR IS THE REPRESSOR OF THE TETRACYCLINE RESISTANCE  
 CC ELEMENT; ITS AMINO-TERMINAL REGION FORMS A HELIX-TURN-HELIX  
 CC STRUCTURE AND BINDS DNA. BINDING OF TETRACYCLINE TO TETR REDUCES  
 CC THE REPRESSOR AFFINITY FOR THE TETRACYCLINE RESISTANCE GENE  
 CC (TETA) PROMOTER OPERATOR SITES.  
 CC -!- INDUCTION: BY THE [MG-TETRACYCLINE]+ COMPLEX.  
 CC -!- SIMILARITY: BELONGS TO THE TETR/ACRR FAMILY OF TRANSCRIPTIONAL  
 CC REGULATORS.  
 DR EMBL; X61367; G48198; -.  
 DR EMBL; X75761; G415985; -.  
 DR EMBL; X00006; G42509; -.  
 DR PIR; A03574; RPECRI.  
 DR PROSITE; PS01081; HTH\_TETR\_FAMILY; 1.  
 KW TRANSCRIPTION REGULATION; REPRESSOR; TRANSDUCIBLE ELEMENT;  
 KW DNA-BINDING; ANTIBIOTIC RESISTANCE; PLASMID.  
 FT DNA\_BIND 26 45 H-T-H MOTIF.  
 FT SITE 54 64 INVOLVED IN BINDING TO [MG-TC]+ (BY  
 FT SIMILARITY).  
 FT METAL 100 100 MAGNESIUM (OF [MG-TC]+ COMPLEX) (BY  
 FT SIMILARITY).  
 FT CONFLICT 65 66 TH -> ST (IN REF. 3).  
 FT CONFLICT 80 80 I -> T (IN REF. 3).  
 FT CONFLICT 154 155 DA -> ES (IN REF. 3).  
 SQ SEQUENCE 216 AA; 23320 MW; F72BDFB1 CRC32;

Query Match 46.8%; Score 505.5; DB 1; Length 216;  
 Best Local Similarity 51.5%; Pred. No. 1.2e-40;  
 Matches 104; Conservative 35; Mismatches 60; Indels 3; Gaps 2;

QY 1 MSRLDKSVINSALLELVGIEGLTTRKLAQKLGVEQPTLYWHVKNKRALDLALAIEML 60  
 Db 1 MTKLPQNTVIRALDILLNVGVNGLTRKLAERLGVQQPALLYWHFKNKRALDLALAEAML 60

QY 61 DRHTHFCPLKGSWQDFLNKAKSFRCALLSHRNKAKVHSDTRPTEKQYETLENQALFL 120  
 Db 61 AENHSHSVPRADDWRSFLIGNARSPQALLAVRDGARIHAGTRPGAPQWETADQALRFL 120

QY 121 COGFSLENALYALSVAHGFTLGCVLDEQHQVAKEREETPTDMP--PLLRQAIELFD 178  
 Db 121 CEAGFSAGDAVNALMTISFTVGAVLEEQAGSDAGER--GGTVEQAPLSPLLRAAIDAFD 179

QY 179 HQGAEPALFLGLELIICGLEKQ 200  
 Db 180 EAGPDAAFEGGLAVIVDGLAKR 201

RESULT 7  
 TERS\_ECOLI  
 ID TERS\_ECOLI STANDARD; PRT; 211 AA.  
 AC P21337;  
 DT 01-MAY-1991 (REL. 18, CREATED)  
 DT 01-MAY-1991 (REL. 18, LAST SEQUENCE UPDATE)  
 DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
 DE TETRACYCLINE REPRESSOR PROTEIN CLASS E.  
 GN TETR.  
 OS ESCHERICHIA COLI.  
 OG PLASMID PSL1503.  
 OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;  
 OC ENTEROBACTERIACEAE.  
 RN [1]  
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-21.  
 RX MEDLINE; 89201249.  
 RA TOVAR K., ERNST A., HILLEN W.;  
 RL MOL. GEN. GENET. 215:76-80(1988).  
 CC -!- FUNCTION: TETR IS THE REPRESSOR OF THE TETRACYCLINE RESISTANCE  
 CC ELEMENT; ITS AMINO-TERMINAL REGION FORMS A HELIX-TURN-HELIX  
 CC STRUCTURE AND BINDS DNA. BINDING OF TETRACYCLINE TO TETR REDUCES  
 CC THE REPRESSOR AFFINITY FOR THE TETRACYCLINE RESISTANCE GENE  
 CC (TETA) PROMOTER OPERATOR SITES.  
 CC -!- SIMILARITY: BELONGS TO THE TETR/ACRR FAMILY OF TRANSCRIPTIONAL

CC REGULATORS.  
 DR EMBL; M34933; G155020; ALT\_SEQ.  
 DR EMBL; X14035; G43056; -.  
 DR PROSITE; PS01081; HTH\_TETR\_FAMILY; 1.  
 KW TRANSCRIPTION REGULATION; REPRESSOR; DNA-BINDING;  
 KW ANTIBIOTIC RESISTANCE; PLASMID.  
 FT DNA\_BIND 26 45 H-T-H MOTIF (POTENTIAL).  
 FT SITE 54 64 INVOLVED IN BINDING TO [MG-TC]+ (BY  
 FT SIMILARITY).  
 FT METAL 100 100 MAGNESIUM (OF [MG-TC]+ COMPLEX) (BY  
 FT SIMILARITY).  
 FT CONFLICT 45 45 V -> L (IN G43056).  
 SQ SEQUENCE 211 AA; 23585 MW; F538C813 CRC32;

Query Match 45.7%; Score 493.5; DB 1; Length 211;  
 Best Local Similarity 50.2%; Pred. No. 1.6e-39;  
 Matches 101; Conservative 29; Mismatches 68; Indels 3; Gaps 1;

QY 1 MSRLDKSVINSALLELVGIEGLTTRKLAQKLGVEQPTLYWHVKNKRALDLALAIEML 60  
 Db 1 MARSLDDVISMALTLTLDSEGLEGLTTRKLAQSLKIQPTLYWHVRNKQTLMMWLSEAIL 60

QY 61 DRHTHFCPLKGSWQDFLNKAKSFRCALLSHRNKAKVHSDTRPTEKQYETLENQALFL 120  
 Db 61 AKHTRSAPLPTESWQQFLQENALSFKALLVHRDGARLHIGTSPTPPQPEQAQALRCL 120

QY 121 COGFSLENALYALSVAHGFTLGCVLDEQHQVAKEREETPTDMPPLLRQAIELFDHQ 180  
 Db 121 CDAGFSVEEALFILOSISHTLGALEE---QATNQIENHNHVIDAAPPLLQAEAFNIQART 177

QY 181 GAEPALFLGLELIICGLEKOL 201  
 Db 178 SAEMAFHGLKSLIFGFSQAL 198

RESULT 8  
 TER7\_VIBAN  
 ID TER7\_VIBAN STANDARD; PRT; 210 AA.  
 AC P51560;  
 DT 01-OCT-1996 (REL. 34, CREATED)  
 DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)  
 DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)  
 DE TETRACYCLINE REPRESSOR PROTEIN CLASS G.  
 GN TETR.  
 OS VIBRIO ANGUILLARUM.  
 OG PLASMID PJA8122.  
 OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;  
 OC VIBRIONACEAE.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=M28122;  
 RX MEDLINE; 93125225.  
 RA ZHAO J., AOKI T.;  
 RL MICROBIOL. IMMUNOL. 36:1051-1060(1992).  
 CC -!- FUNCTION: TETR IS THE REPRESSOR OF THE TETRACYCLINE RESISTANCE  
 CC ELEMENT; ITS AMINO-TERMINAL REGION FORMS A HELIX-TURN-HELIX  
 CC STRUCTURE AND BINDS DNA. BINDING OF TETRACYCLINE TO TETR REDUCES  
 CC THE REPRESSOR AFFINITY FOR THE TETRACYCLINE RESISTANCE GENE  
 CC (TETA) PROMOTER OPERATOR SITES.  
 CC -!- INDUCTION: BY THE [MG-TETRACYCLINE]+ COMPLEX.  
 CC -!- SIMILARITY: BELONGS TO THE TETR/ACRR FAMILY OF TRANSCRIPTIONAL  
 CC REGULATORS.  
 DR EMBL; S52438; G262929; -.  
 DR PROSITE; PS01081; HTH\_TETR\_FAMILY; 1.  
 KW TRANSCRIPTION REGULATION; REPRESSOR; DNA-BINDING;  
 KW ANTIBIOTIC RESISTANCE; PLASMID.  
 FT DNA\_BIND 26 45 H-T-H MOTIF (POTENTIAL).  
 FT SITE 54 64 INVOLVED IN BINDING TO [MG-TC]+ (BY  
 FT SIMILARITY).  
 FT METAL 100 100 MAGNESIUM (OF [MG-TC]+ COMPLEX) (BY  
 FT SIMILARITY).  
 SQ SEQUENCE 210 AA; 23595 MW; 811CB332 CRC32;



```
Query Match      43.8%; Score 473; DB 1; Length 210;
Best Local Similarity 47.3%; Pred. No. 1.3e-37;
Matches 95; Conservative 32; Mismatches 72; Indels 2; Gaps 1;

QY 1 MSRLDKSVINSALELLNEVGIEGLTRKLAQKLGVEOPTLYWHVKNKRALLDALAIEML 60
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1 MYKLDKGTIVIAAGLELLNEVGMDSLTRKLAERLKVQOPALYWHFQNKRALLDALPEAML 60

QY 61 DRHHTFCPLKGSWODFLRNKAKSFRCALLSHRNKAKVHSDTRTEKQYETLENQLAFL 120
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 61 RERHTRSLPEENEDWVFLKNALSFRTALLSYRGARIHAGTRTEPNFGTAETQIRFL 120

QY 121 COGFSLNALYALSVAHGHTLGCVLEDOHQVAKERTPTTDSMP--LLRQAIELFD 178
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 121 CAEGFCPKAVWALRAVSHVVGSLVEQASDADERVPDPVSEQAPSFLLHVFHELE 180

QY 179 HOGAEPAPFLFGLLEIICGLEK 199
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 181 TDGMDAAFNGLDLSIAGFER 201

RESULT 9
Y07H_MYCTU      STANDARD;          PRT;    261 AA.
AC Q11023;
DT 01-OCT-1996 (REL. 34, CREATED)
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DE 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL TRANSCRIPTIONAL REGULATOR CY02B10.17C.
GN MYCY02B10.17C.
OS MYCOBACTERIUM TUBERCULOSIS.
OC PROKARYOTA; FIRMICUTES; ACTINOMYCETALES; MYCOBACTERIACEAE.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-H37V;
RA MURPHY L., HARRIS D., BARRELL B.G., RAJANDREAM M.A., WALSH S.V.;
RL SUBMITTED (JUN-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -!- SIMILARITY: BELONGS TO THE TETR/ACRR FAMILY OF TRANSCRIPTIONAL
   REGULATORS.
DR EMBL; 275555; 250355; -.
KW HYPOTHETICAL PROTEIN; TRANSCRIPTION REGULATION; DNA-BINDING.
FT DNA_BIND 38 57 H-T-H MOTIF (POTENTIAL).
SQ SEQUENCE 261 AA; 28253 MW; 8CBCC98 CRC32;

Query Match      9.9%; Score 106.5; DB 1; Length 210;
Best Local Similarity 26.8%; Pred. No. 0.0041;
Matches 51; Conservative 23; Mismatches 73; Indels 43; Gaps 8;

QY 6 KSKVINSALELLNEVGIEGLTRKLAQKLGVEOPTLYWHVKNKRALLDALAIEMLDHRHT 65
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 12 KEHMLAALETFYRKGIARTSLNEIAAGVTRGALYWHFKNKEDLFDALFORICDDIEN 71

QY 66 ---HFCPLKGSWODFLRNKAKSFRCALLSHRNKAKVHSDTRTEKQYETLENQLAFLC 121
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 72 CIAQDAADAEAGSWTVF-RHLLHFFERLQSDNDIHYKFH-----NILFLKC 116

QY 122 QOGFSLNALYALSVAHGHTLGCVLEDOHQVAKERTPTTDSMPPLLRQAIIE---LFD 178
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 117 EH--TEQNA--AVIAIA-----RKQAIWREKIT-----AVLTEAVENQDLAD 155

QY 179 HOGAEPAPFLF 188
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 156 DLDKETAVIF 165

RESULT 11
YCFQ_ECOLI      STANDARD;          PRT;    236 AA.
AC P75952;
DT 15-JUL-1998 (REL. 36, CREATED)
DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL TRANSCRIPTIONAL REGULATOR IN NDH-MFD INTERGENIC REGION.
GN YCFQ.
OS ESCHERICHIA COLI.
OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;
OC ENTEROBACTERIACEAE.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-K12 / MG1655;
RA BLATTNER F.R., PLUNKETT G. III, MAYHEW G.F., PERNA N.T., GLASNER F.D.;
RL SUBMITTED (JAN-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -!- SIMILARITY: BELONGS TO THE TETR/ACRR FAMILY OF TRANSCRIPTIONAL
   REGULATORS.
DR EMBL; AE000211; G1787354; -.
KW ECOCENE; EGI3435; YCFQ.
DR PROSITE; PS01081; HTH_TETR_FAMILY; FALSE_NEG.
KW HYPOTHETICAL PROTEIN; TRANSCRIPTION REGULATION; REPRESSOR;
   DNA-BINDING.
```

```
Query Match      15.6%; Score 168.5; DB 1; Length 261;
Best Local Similarity 27.0%; Pred. No. 7.8e-09;
Matches 61; Conservative 32; Mismatches 78; Indels 55; Gaps 6;

QY 4 LDKSVINSALELLNEVGIEGLTRKLAQKLGVEOPTLYWHVKNKRALLDALAIEMLDHRH 63
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 16 INPEDIISGAFELAAQVSDINLWPLGLKHLGVGTYSIYVFRKDDLLNAMTDRALSKY 75

QY 64 HTHFCPLKGSWODFLRNKAKSFR-----CALL-----SHRNKAKVHSDTRP 105
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 76 VFATPIEAGDWRETLRNHARSKRMTFADNPVLCDLILIRAALSPKTLRGA----- 128

QY 106 TEKQYETLENQLAFLCOGFSLNALYALSVAHGHTLGCVLEDO-----E 150
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 128 -----QEMEKATIANLVTAGLSLEDAFDIYSVSVHVRGVSVDRLSRKKSQAGSGPSAIE 182

QY 151 HOVAKEERTPTTDSMPPLLRQAIELFDHQA--EPAPFLFGLLELII 194
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 183 HPVADIPATT-----PILAHATGRGHRICGAPDETNEYGLECIL 221

RESULT 10
MTRR_NEIGO
ID MTRR_NEIGO      STANDARD;          PRT;    210 AA.
AC P39897;
```

FT DNA\_BIND 67 86 H-T-H MOTIF (POTENTIAL).  
SQ SEQUENCE 236 AA; 26204 MW; AD1F64FC CRC32;

Query Match 8.6%; Score 92.5; DB 1; Length 236;  
Best Local Similarity 25.2%; Pred. No. 0.098;  
Matches 35; Conservative 18; Mismatches 63; Indels 23; Gaps 5;

QY 5 DSKVINSALELNEVGIEGLTRKLAQKLGVEOPTLYHVHVKNRALLDALAIEMLDRL-- 63  
DB 46 DRDAALDKAMFLWQHOYEATSLADLVATGCAKAPTLYAEFTNKGLFRAVLDRYIDREA 105  
QY 63 --HHTH-FCPLKG--ESWQDFLRNKASF-----RCALLSHENGAKVHSDTR 104  
DB 106 AKHEAQLFCBEKSVESALADYFAAIANCFTSKDTPAGCFMINNCTTSPDGGDIANTLKS 165  
QY 105 PTEKQYETLENQALFLCQQ 123  
DB 166 RHAMQERTLQO---FLCQR 181

RESULT 12  
UIUR\_ECOLI STANDARD; PRT; 196 AA.  
AC Q39431;  
DT 01-NOV-1997 (REL. 35, CREATED)  
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)  
DE UID OPERON REPRESSOR (GUS OPERON REPRESSOR).  
GN UIDR OR GUSR.  
OS ESCHERICHIA COLI.  
OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;  
OC ENTEROBACTERIACEAE.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-K12;  
RA WILSON K.J.; JEFFERSON R.A.;  
RL SUBMITTED (OCT-1994) TO EMBL/GENBANK/DBJ DATA BANKS.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-K12 / MG1655;  
RA BLATTNER F.R.; PLUNKETT G. III; MAYHEW G.F.; PERNA N.T.; GLASNER F.D.;  
RL SUBMITTED (JAN-1997) TO EMBL/GENBANK/DBJ DATA BANKS.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN-K12;  
RA AIBA H.; BABA T.; FUJITA K.; HAYASHI K.; HONJO A.; HORIIUCHI T.;  
RA IKEMOTO K.; INADA T.; ISONO K.; ISONO S.; ITOH T.; KANAI K.; KASAI H.;  
RA KASHIMOTO K.; KIM S.; KIMURA S.; KITAGAWA M.; KITAKAWA K.; MAKINO K.;  
RA MASUDA S.; MIKI T.; MIZOBUCHI K.; MORI H.; MOTOMURA K.; NAKAMURA Y.;  
RA NASHIMOTO H.; NISHIO Y.; OSHIMA T.; SAITO N.; SAMEI G.; SEKI Y.;  
RA TAGAMI H.; TAKEMOTO K.; WADA C.; YAMAMOTO Y.; YANO M.;  
RL SUBMITTED (DEC-1996) TO EMBL/GENBANK/DBJ DATA BANKS.  
CC -!- FUNCTION: REPRESSOR FOR THE UIDRABC (GUSRABC) OPERON.  
CC -!- SIMILARITY: BELONGS TO THE TETR/ACRR FAMILY OF TRANSCRIPTIONAL  
CC REGULATORS.  
DR EMBL; M14641; G868019; -;  
DR EMBL; AE000257; G1787904; -;  
DR EMBL; D90805; G1742672; -;  
DR ECOGENE; EG12667; UIDR.  
DR PROSITE; PS01081; HTH TETR FAMILY; 1.  
KW TRANSCRIPTION REGULATION; DNA-BINDING; REPRESSOR.  
FT DNA\_BIND 33 52 H-T-H MOTIF (POTENTIAL).  
SQ SEQUENCE 196 AA; 21799 MW; 6C1840FC CRC32;

Query Match 8.5%; Score 91.5; DB 1; Length 196;  
Best Local Similarity 24.0%; Pred. No. 0.097;  
Matches 18; Conservative 24; Mismatches 32; Indels 1; Gaps 1;

QY 6 KSKVINSALELNEVGIEGLTRKLAQKLGVEOPTLYHVHVKNRALLDALAIEMLDRLHHT 65  
DB 13 RTRILNAAREIFSENGFHSASMAKICKSCAISPGLYHHFISKEALLQAIILQDOERALA 72

QY 66 HFC-PLKGESWQDFL 79  
DB 73 RREPPIEGIHFDYM 87

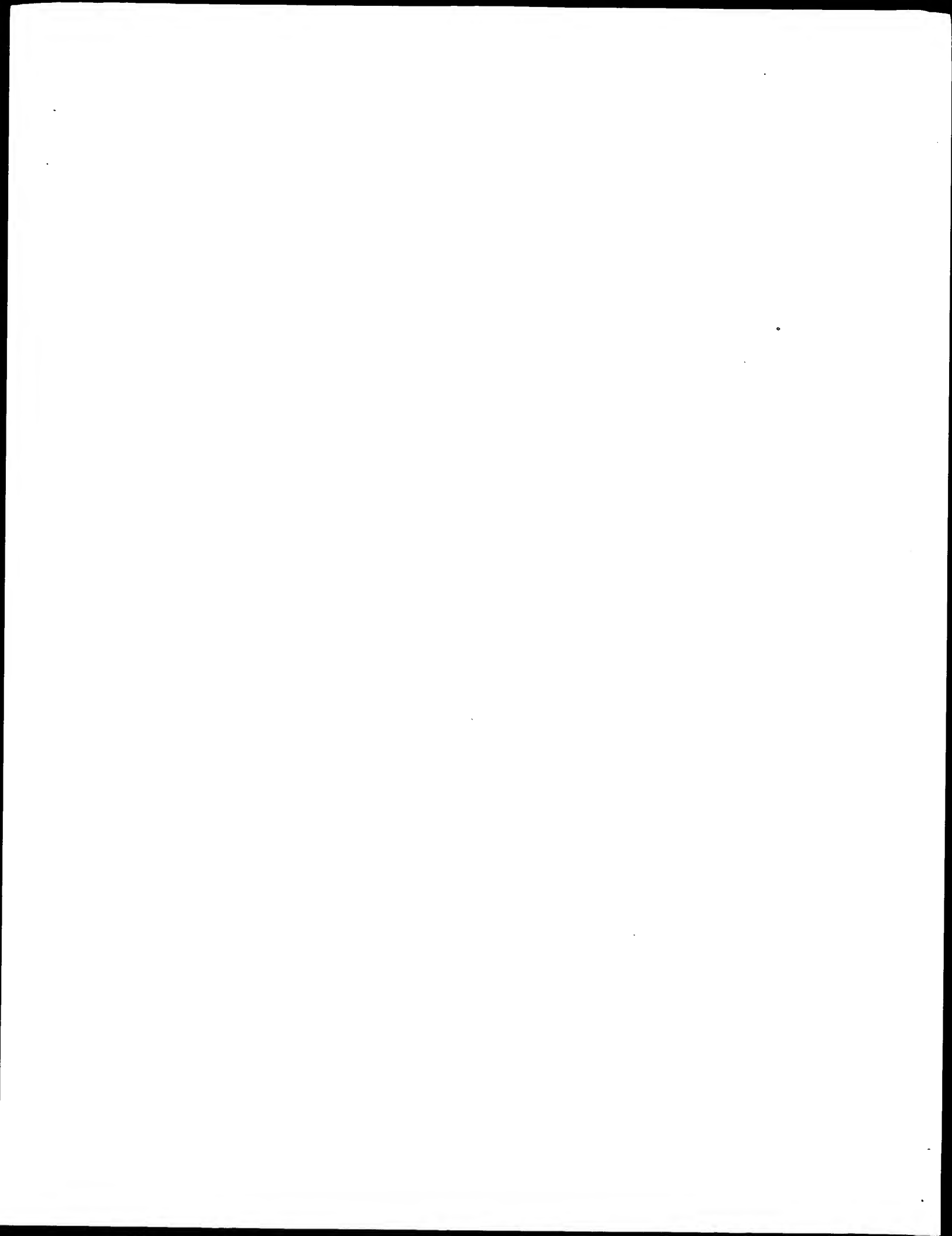
RESULT 13  
PHSL\_IPOBA STANDARD; PRT; 955 AA.  
ID PHSL\_IPOBA  
AC P27598;  
DT 01-AUG-1992 (REL. 23, CREATED)  
DT 01-AUG-1992 (REL. 23, LAST SEQUENCE UPDATE)  
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
DE ALPHA-GLUCAN PHOSPHORYLASE, L ISOZYME PRECURSOR (EC 2.4.1.1)  
DE (STARCH PHOSPHORYLASE L).  
OS IPOMOEA BATATAS (SWEET POTATO) (BATATE).  
OC EUKARYOTA; PLANTA; EMERYOPHYTA; ANGIOSPERMAE; DICOTYLEDONEAE;  
OC SOLANALES; CONVULVACEAE.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA LIN C.T.; YEH K.W.; LEE P.D.; SU J.C.;  
RL PLANT PHYSIOL. 0:0-0(1992).  
CC -!- FUNCTION: PHOSPHORYLASE IS AN IMPORTANT ALLOSTERIC ENZYME IN  
CC CARBOHYDRATE METABOLISM. ENZYMES FROM DIFFERENT SOURCES DIFFER IN  
CC THEIR REGULATORY MECHANISMS AND IN THEIR NATURAL SUBSTRATES.  
CC HOWEVER, ALL KNOWN PHOSPHORYLASES SHARE CATALYTIC AND STRUCTURAL  
CC PROPERTIES.  
CC -!- CATALYTIC ACTIVITY: (1,4-ALPHA-D-GLUCOSYL)(N) + ORTHOPHOSPHATE =  
CC (1,4-ALPHA-D-GLUCOSYL)(N-1) + ALPHA-D-GLUCOSE 1-PHOSPHATE.  
CC -!- COFACTOR: PYRIDOXAL PHOSPHATE.  
CC -!- SUBCELLULAR LOCATION: CHLOROPLAST; AMYLOPLAST.  
CC -!- SIMILARITY: BELONGS TO THE GLYCOGEN PHOSPHORYLASE FAMILY.  
DR EMBL; M64362; G168276; -;  
DR HSP; P00489; 9GPB.  
DR PROSITE; PS0102; PHOSPHORYLASE; 1.  
DR TRANSFERASE; GLYCOSYLTRANSFERASE; CARBOHYDRATE METABOLISM;  
KW ALLOSTERIC ENZYME; PYRIDOXAL PHOSPHATE; TRANSIT PEPTIDE; CHLOROPLAST;  
KW AMYLOPLAST; MULTIGENE FAMILY.  
FT TRANSIT 1 43 CHLOROPLAST (POTENTIAL).  
FT CHAIN 44 955 ALPHA-GLUCAN PHOSPHORYLASE, L ISOZYME.  
FT BINDING 801 801 PYRIDOXAL PHOSPHATE (BY SIMILARITY).  
SQ SEQUENCE 955 AA; 108520 MW; 86E3D50F CRC32;

Query Match 8.0%; Score 86; DB 1; Length 955;  
Best Local Similarity 30.1%; Pred. No. 2.2;  
Matches 56; Conservative 19; Mismatches 67; Indels 44; Gaps 10;

QY 1 MSRL-----DKSVINSALELNEVGIEGL--TTRKLAQKLGVEOPTLYHVH---- 46  
DB 1 MSRLSGITPRADRSQFQNPRLIAVDPRTAGLQRTKTLVKCVLDETQTIQHVYTE 60  
QY 46 KNRALLDALAIEMLDRLHHTFCPLKGESWQDFLRNK-----AKSFRCALLSHRNGAKVH 100  
DB 61 KNEGTLLDAASIASIKYHAEFSP--AFSPERFELPKAYFATAQSVRDALIVNNA---- 115  
QY 101 SDTRPTEKQYETLENQALFLCQ-----QGFSLNAYALVAVGHFT-----LGCVLEDOEH 151  
DB 115 -----TYDYEXLNMKQAYLSMEFLQGRALLNAGLNLGTGEYAEALNKLGNLEN--- 167  
QY 152 QVAKEE 157  
DB 167 -VASKE 171

RESULT 14  
THTL\_SCHPO STANDARD; PRT; 577 AA.  
ID THTL\_SCHPO  
AC Q09684;  
DT 01-NOV-1995 (REL. 32, CREATED)  
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)  
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
DE NUCLEAR FUSION PROTEIN THTL.

Search completed: June 9, 1999, 13:01:03  
Job time: 9122 sec



GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 9, 1999, 13:00:33 ; Search time 28.57 Seconds  
(without alignments)  
399.721 Million cell updates/sec

Title: US-08-486-814-19

Perfect score: 1080

Sequence: 1 MSRLDKSVINSALLELNEV.....FGLELIICGLEKQLKCESGS 207

Scoring table: BLOSUM62

Searched: 180763 seqs, 55169189 residues

Database : SPTRMBL\_8.\*

1: sp\_fungi.\*  
2: sp\_human.\*  
3: sp\_invertebrate.\*  
4: sp\_mammal.\*  
5: sp\_mhc.\*  
6: sp\_organelle.\*  
7: sp\_phage.\*  
8: sp\_plant.\*  
9: sp\_bacteria.\*  
10: sp\_rodent.\*  
11: sp\_virus.\*  
12: sp\_invertebrate.\*  
13: sp\_unclassified.\*  
14: sp\_archaea.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	597	55.3	126	9	Q59466 haemophilus
2	147.5	13.7	259	9	Q53901 streptomyce
3	124.5	11.5	242	9	O52558 amycolatops
4	123	11.4	191	9	O34892 bacillus su
5	105.5	9.8	210	9	O59608 neisseria g
6	103.5	9.6	192	9	O67927 aquifex aeo
7	102	9.4	189	9	O34843 bacillus su
8	99	9.2	196	9	O53789 mycobacteri
9	98.5	9.1	186	9	O51597 pseudomonas
10	96.5	8.9	246	9	O53295 mycobacteri
11	94	8.7	179	9	O67157 aquifex aeo
12	93.5	8.7	236	9	P73186 synchocyst
13	93	8.6	217	9	O87854 streptomyce
14	91.5	8.5	213	9	O53310 mycobacteri
15	91.5	8.5	194	9	O07001 bacillus su
16	89	8.2	202	9	O51730 pseudomonas
17	88.5	8.2	2472	10	P70477 rattus norv
18	88.5	8.2	2472	10	O88663 rattus norv
19	86.5	8.0	2477	2	Q13186 homo sapien
20	86.5	8.0	435	2	O75901 staphylococ
21	86.5	8.0	185	9	O70020 staphylococ
22	86	8.0	340	8	O42863 ipomoea bat
23	86	8.0	488	9	O54830 streptococ
24	86	8.0	581	9	O54835 streptococ
25	86	8.0	205	9	Q51516 pseudomonas
26	84	7.8	539	9	O54719 streptococ
27	83.5	7.7	216	9	P95100 mycobacteri
28	83	7.7	2748	1	O03767 saccharomyc
29	83	7.7	198	9	P96676 bacillus su

Q59802 staphylococ  
Q29994 archaeoglob  
O18399 drosophila  
Q91290 pleurodeles  
O34970 bacillus su  
O19126 macaca fasc  
O19127 macaca mula  
O25951 helicobacte  
O86852 streptomyce  
O53165 mycobacteri  
O07388 mycobacteri  
O90631 gallus gall  
O59306 clostridium  
P94548 bacillus su  
O15087 homo sapien  
P79138 cercopithec

#### ALIGNMENTS

##### RESULT 1

Q59466  
ID O59466 PRELIMINARY; PRT; 126 AA.  
AC O59466;  
DT 01-NOV-1996 (TREMREL. 01, CREATED)  
DT 01-NOV-1996 (TREMREL. 01, LAST SEQUENCE UPDATE)  
DT 01-NOV-1998 (TREMREL. 08, LAST ANNOTATION UPDATE)  
DE 30S RIBOSOMAL PROTEIN S21.  
OS HAEMOPHILUS PARAINFLUENZA.  
OC BACTERIA; PROTEOBACTERIA; GAMMA SUBDIVISION; PASTEURELLACEAE;  
OC HAEMOPHILUS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 87137315.  
RA HEUER C., HICKMAN R.K., CURIALE M.S., HILLEN W., LEVY S.B.;  
RT "Constitutive expression of tetracycline resistance mediated by a  
RT Tn10-like element in Haemophilus parainfluenzae results from a  
RT mutation in the repressor gene."  
RL J. BACTERIOL. 169:990-994(1987).  
CC -!- SIMILARITY: BELONGS TO THE S21P FAMILY OF RIBOSOMAL PROTEINS.  
DR EMBL; M15539; G148989;  
DR PROSITE; PS01081; HTH\_TETR\_FAMILY; 1.  
DR PFAM; PF00440; tetr; 1.  
KW RIBOSOMAL PROTEIN.  
SQ SEQUENCE 126 AA; 14541 MW; 107F233B CRC32;

Query Match 55.3%; Score 597; DB 9; Length 126;  
Best Local Similarity 92.0%; Pred. No. 1.3e-45;  
Matches 115; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 1 MSRLDKSVINSALLELNEVIEGLTRKLAQKLGVEQPTLYWHVKNRALLDALAIEML 60  
|||||

Db 1 MSRLDKSVINSALLELNEVIEGLTRKLAQKLGVEQPTLYWHVKNRALLDALAIEML 60  
|||||

QY 61 DRHHTHFCPLKESQWDFLRNKAQKFRCALLSHRNKAVHSDTRTEKQYETLENQALFL 120  
|||||

Db 61 DRHHTHFCPLKESQWDFLRNKAQKFRCALLSHRNKAVHSDTRTEKQYETLENQALFL 120  
|||||

QY 121 COQGF 125  
:|

Db 121 ANKVF 125  
:|

##### RESULT 2

Q53901 PRELIMINARY; PRT; 259 AA.

ID Q53901

AC Q53901;

DT 01-NOV-1996 (TREMREL. 01, CREATED)

DT 01-NOV-1996 (TREMREL. 01, LAST SEQUENCE UPDATE)

DT 01-NOV-1998 (TREMREL. 08, LAST ANNOTATION UPDATE)

DE ORF1-4 (ACT11).

RESULT	4
C34892	
ID	034892 PRELIMINARY; PRT; 191 AA.
AC	034892;
DT	01-JAN-1998 (TREMREL. 05, CREATED)
DT	01-JAN-1998 (TREMREL. 05, LAST SEQUENCE UPDATE)
DT	01-NOV-1998 (TREMREL. 08, LAST ANNOTATION UPDATE)
DE	TRANSCRIPTION REGULATOR.
GN	YOB.
OS	BACILLUS SUBTILIS.
OC	BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; BACILLACEAE;
CC	BACILLUS.
RN	[1]
RP	SEQUENCE FROM N.A.
RA	LAPIDUS A., GALLERON N., SOROKIN A., EHRLICH D.;
RL	SUBMITTED (NOV-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
RN	[2]
RP	SEQUENCE FROM N.A.
RC	STRAIN=168;
RX	MEDLINE; 98044033.
RA	KUNST F., OGASAWARA N., MOSZER I., ALBERTINI A.M., ALLONI G.,
RA	AZEVEDO V., BERTERO M.G., BESSIERES P., BOLOTIN A., BORCHERT S.,
RA	BORRIS R., BOURSIER L., BRANS A., BRAUN M., BRIGHELL S.C., BRON S.,
RA	BROUILLET S., BRUSCHI C.V., CALDWELL B., CAPUANO V., CARTER N.M.,
RA	CHOI S.K., CODANI J.J., CONNERTON I.F., CUMMINGS N.J., DANIEL R.A.,
RA	DENIZOT F., DEVINE K.M., DUSTERHOFT A., EHRLICH S.D., EMMERSON P.T.,
RA	ENTIAN K.D., ERINGTON J., FABRET C., FERRARI E., FOULGER D., FRITZ C.
RA	FUJITA M., FUJITA Y., FUNA S., GALIZZI A., GALLERON N., GHIM S.Y.,
RA	GLASER P., GOFFEAU A., GOLIGHTLY E.J., GRANDI G., GUISEPI G.,
RA	GUY B.J., HAGA K., HAIECH J., HARWOOD C.R., HENAUT A., HILBERT H.,
RA	HOLSAPPEL S., HOSONO S., HULLO M.F., ITAYA M., JONES L., JORIS B.,
RA	KARAMATA D., KASHARA Y., KLAEPR-BLANCHARD M., KLEIN C., KOBYAVASHI Y.,
RA	KOEITER P., KONINGSTEIN G., KROGH S., KUENO M., KURITA K., LAPIDUS A.
RA	LARDINOIS S., LAUBER J., LAZAREVIC V., LEW S.M., LEVINE A., LIU H.
RA	MASUDA S., MAUEL C., MEDIGUE C., MEDINA N., MELLADO R.P., MIZUNO M.,
RA	MOESTL D., NAKAI S., NOBACK M., NOONE D., O'REILLY M., OGAWA K.,
RA	OGIWARA A., OUDEGA B., PARK S.H., PARRO V., POHL T.M., PORTELETTE D.,
RA	PORWOLIK S., PRESCOTT A.M., PRESSECAN E., PUJIC P., PURNELLE B.,
RA	RAPOPORT G., REY M., REYNOLDS S., RIEGER M., RIVOLTA C., ROCHA E.,
RA	ROCHE B., ROSE M., SADATE Y., SATO T., SCANLAN E., SCHLEICH S.,
RA	SCHROETER R., SCOFFONE F., SEKIGUCHI J., SEKORSKA A., SEROR S.J.,
RA	SEROR P., SHIN B.S., SOLDÓ B., SOROKIN A., TACCIONI E., TAKAGI T.,
RA	TAKAHASHI H., TAKEMARU K., TAKEUCHI M., TAMAKOSHI A., TANAKA T.,
RA	TERESTRA P., TEGNONI A., TOSATO V., UCHIYAMA S., VANDENBOL M.,
RA	VANNIER F., VASSAROTTI A., VIARI A., WAMBUUT R., WEDLER E., WEDLER H.,
RA	WEITZENEGGER T., WINTERS P., WIPAT A., YAMAMOTO H., YAMANE K.,
RA	YASUDOMOTO K., YATA K., YOSHIDA K., YOSHIKAWA H.F., ZUMSTEIN E.,
RA	YOSHIKAWA H., DANCHIN A.;
RT	"The complete genome sequence of the gram-positive bacterium Bacillus
RT	subtilis.;"
RL	NATURE 390:249-256(1997).
RN	[3]
RP	SEQUENCE FROM N.A.

[illegible]

[illegible]



Query Match 8.9%; Score 96.5; DB 9; Length 246;  
Best Local Similarity 18.5%; Pred. No. 0.29;  
Matches 40; Conservative 45; Mismatches 80; Indels 51; Gaps 6;  
Qy 6 KSKVNSALELLNEVGIETGLTRTKLAQKLGVEQPTLYWHVKNKRALLDALALEMLDRHHT 65

```

Query Match      8.7%: Score 94; DB 9; Length 179;
Best Local Similarity 21.4%; Pred. No. 0.33;
Matches         47; Conservative 33; Mismatches 68; Indels 72; Gaps 9;

QY   6 KSKVNSALELLNEVGIEGLITRKLAKLGVBPOTLYHVHNKRALLDALATEMLDRHHT 65
     | : : | | : | : : : | : | : | : | : | : | : | : | : | : | : |
DB   4 KERILLEVKSELPFEKGYOQTSEEIVKRNLSKGAFYHFHSKEELI---TEIIERTHK 59
     | : : | | : | : : : | : | : | : | : | : | : | : | : | : | : |

QY   66 HFCPLKGESWQDFLNKAASFRCALLSHRNGAKVHSIDFRPTEKYETL--ENQLAF---- 120
     | | | | : | | | : | : | : | : | : | : | : | : | : | : | : |
DB   60 KIISL-----FEENKETPEELL-----EMFLEVLYREKKVVIFLF 96
     | : : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

QY   120 ---LCOOGF-----SLENALYAL-----SAVGHTFLGCVLEDPQOEQAVERE 159
     || : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB   97 DILCSEKFNIYFEKIEDAKRFEFLEKHFFSKAEIISEILGFLROLILHAVIKEERE 156
     | : : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

QY   160 TPTTDSMPPLLRLQATLEFDHOGAEP AFLGLELIICGLEK 199
     | : : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB   157 LP---FLKEKLE-----GLKLIFEKVKK 177
     | : : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

RESULT 12
P73186 PRELIMINARY; PRT; 236 AA.
ID P73186
AC P73186;
DT 01-FEB-1997 (TREMBREL. 02, CREATED)

```

```

RT      "A set of ordered cosmids and a detailed genetic and physical map for
RT      the 8 Mb Streptomyces coelicolor A3(2) chromosome." ;
RL      MOL. MICROBIOL. 21:77-96(1996).
DR      EMBL; AL031013; E1309426; -.
SQ      SEQUENCE      217 AA;  23650 MW;  FBPEC38D CRC32;

      Query Match      8.6%;  Score 93;  DB 9;  Length 217;
      Best Local Similarity 37.9%;  Pred. No. 0.51;
      Matches 22;  Conservative 11;  Mismatches 25;  Indels 0;  Gaps

QY      3 RLDKSKVINSALELLNEVGIEGTTTRKLAQKLGVEQPTLYWHVKNKRALDMLAEML 60
      | : : : : : | | | | : : : : : | | : : | | | | | : | : | : |
DB      6 RWSTEQILDAAELLLAGDAFTSVRKLAASLGTDSSSLYRHFERNKTELLRAVADRIL 63

RESULT 14
053310      PRELIMINARY;      PRT;      213 AA.
AC      053310;
DT      01-JUN-1998 (TREMBLREL. 06, CREATED)
DT      01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)
DT      01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE      HYPOTHETICAL 23.1 KD PROTEIN.
GN      MYV014.04C.
GS      MYCOBACTERIUM TUBERCULOSIS.
OS      BACTERIA; FIRMICUTES; ACTINOBACTERIA; MYCOBACTERIA; MYCOBACTERIACEAE;
OC      MYCOBACTERIUM.
[1]
RN      RP
RP      SEQUENCE FROM N.A.
RC      STRAIN-H37RV;
RL      OLIVER K., HARRIS D.;
RL      SUBMITTED (JAN-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
[2]
RN      RP
RP      SEQUENCE FROM N.A.
RC      STRAIN-H37RV;
RL      BARRELL B.G., RAJANDREAM M.A.;
RL      SUBMITTED (JAN-1997) TO EMBL/GENBANK/DBJ DATA BANKS.

```

```

RA COLE S.F.;;
RT "An integrated map of the genome of the tubercle bacillus,
RT Mycobacterium tuberculosis H37Rv, and comparison with Mycobacterium
RT leprae.";
RT PROC. NATL. ACAD. SCI. U.S.A. 93:3132-3137(1996).
DR EMBL; AL021646; E1248751; -.
SQ HYPOTHETICAL PROTEIN.
KW SEQUENCE 213 AA; 23120 MW; 2316B357 CRC32;

Query Match      8.5%; Score 91.5; DB 9; Length 213;
Best Local Similarity 38.9%; Pred. No. 0.67;
Matches 21; Conservative 11; Mismatches 19; Indels 3; Gaps 0

QY      8 KVINSALELLNEVGEGILTRKLAQKLGVEQPTLYMHWKNKRALDA---LATE 58
      :: : ||| : : : : : ||| : : : : : ||| : : : : : ||| :
db     14 RLIGAAELIALRGYSSTSRDIAAAGVGEQPAIYKHFSAKRDIILAAVLRLAVE 67

RESULT 15
007001 PRELIMINARY; PRT; 194 AA.
ID O07001
AC O07001;
DT 01-JUL-1997 (TREMBLREL 04, CREATED)
DI 01-JUL-1997 (TREMBLREL 04, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL 08, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL 22.2 KD PROTEIN.
OS YVDI.
OS BACILLUS SUBTILIS.

```



